

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2003, 15:03:13 ; Search time 3307 Seconds
(without alignments)
9847.608 Million cell updates/sec

Title: US-09-921-992-3

Perfect score: 1119
Sequence: 1 atgcataccagagctccaat.....ttcagcaggttgaataaa 1119

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb_da.*
2: gb_htg.*
3: gb_in.*
4: gb_cm.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: gb_da.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_cm.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vl.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rnd.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1119	100.0	1119	6 AX036302	AX036302 Sequence
2	1119	100.0	1119	6 AX038912	AX038912 Sequence
3	1119	100.0	1119	6 AX050487	AX050487 Sequence
4	1119	100.0	1119	6 AX039343	AX039343 Sequence
5	1119	100.0	1697	1 EC6451	EC6451 E.coli gcpE
6	1119	100.0	13176	1 AE000338	AE000338 Escherich
7	1093.4	97.7	11521	1 AE003515	AE003515 Escherich
8	1093.4	97.7	11521	1 AE005481	AE005481 Escherich
9	1093.4	97.7	296827	1 AP002561	AP002561 Escherich
10	1093.4	90.3	1010	6 AR208328	AR208328 Sequence
11	1010	90.3	1010	6 AR208328	AR208328 Sequence
12	887	79.3	23647	1 AE008814	AE008814 Salmonell
13	885.4	79.1	145050	1 AL627275	AL627275 Salmonell
14	710	63.4	4775	1 D90881	D90881 E.coli geno
15	690	61.7	16214	1 D90880	D90880 E.coli geno
16	685.4	59.4	11824	1 AE004161	AE004161 Mus muscu
17	665	56.4	1345	1 PS067931	PS067931 Vibrio ch
18	631.6	56.4	1345	6 AR208326	AR208326 Sequence
19	631.6	56.4	1345	6 AR208326	AR208326 Sequence
20	631.6	56.4	1345	6 AR208326	AR208326 Sequence
21	622	55.6	10507	1 U32721	U32721 Haemophilus
22	620.6	55.5	11161	1 AE013738	AE013738 Yersinia
23	620.6	55.5	204050	1 AJ414154	AJ414154 Yersinia
24	619.8	55.4	1079	6 AR027992	AR027992 Sequence
25	619.8	55.4	1079	6 AR027992	AR027992 Sequence
26	603.6	53.9	11268	1 AE006237	AE006237 Pasteurel
27	534.6	47.8	886	6 AR027995	AR027995 Sequence
28	534.6	47.8	886	6 AR027995	AR027995 Sequence
29	515.4	46.1	17148	1 AE004798	AE004798 Pseudomon
30	403.6	36.1	347550	1 AP001118	AP001118 Buchnera
31	388.8	34.7	9608	1 AE014104	AE014104 Caulobact
32	382.6	34.3	9999	1 AE005761	AE005761 Homo sapi
33	382.6	34.2	33140	6 AX067448	AX067448 Sequence
34	317.2	28.3	987	6 HSA32568	AJ325686 Homo sapi
35	293.4	26.2	1083	6 AX433277	D29116 Bacillus su
36	277.2	24.8	218470	1 BSUB0013	D84432 Bacillus su
37	277.2	24.8	282700	1 BAC1H642	AR027994 Sequence
38	269	24.0	886	6 AR027994	AR027994 Sequence
39	269	23.6	43632	1 SC6A5	AL048485 Streptomy
40	263.6	23.1	39744	1 SC5H4	AL355913 Streptomy
41	258.2	22.8	6157	6 AX416761	AX416761 Sequence
42	254.6	22.5	1107	6 AX413788	AX413788 Sequence
43	251.4	22.5	1107	6 AX415815	AX415815 Sequence
44	251.4	22.5	240050	1 AL591979	AL591979 Listeria
45	251.4	22.5	240050	1 AL591979	AL591979 Listeria

ALIGNMENTS

RESULT 1
LOCUS AX036302
DEFINITION Sequence 29 from Patent EP1043403.
ACCESSION AX036302
VERSION AX036302.1 GI:11225912
KEYWORDS
SOURCE Escherichia coli.
ORGANISM Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
FEATURES
source 1..1119
/organism="Escherichia coli"
/db_xref="taxon:562"

Pred. No. 19 the number of results predicted by chance to have a

BASE COUNT 272 a 279 c 317 g 251 t

Query Match 100.0%; Score 1119; DB 6; Length 1119;
Best Local Similarity 100.0%; Pred. No. 9.9e-271;
Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGCATACACAGGCTCATATCAAGCTAGAAAAATCAACAGCTATTTACGTTGGGAATGTG 60
 Db 1 ATGCATACACAGGCTCATATCAAGCTAGAAAAATCAACAGCTATTTACGTTGGGAATGTG 60
 OY 61 CCGATTGGCGATGTGCTGCATCGCCGTACAGTCCATACCAATACGGGTACAGACAGAC 120
 Db 61 CCGATTGGCGATGTGCTGCATCGCCGTACAGTCCATACCAATACGGGTACAGACAGAC 120
 OY 121 GTGGAAGCAAGGCTCAATGAATCAAGGGCGTGAAGCGGTTGGCGCTGATATCGCTT 180
 Db 121 GTGGAAGCAAGGCTCAATGAATCAAGGGCGTGAAGCGGTTGGCGCTGATATCGCTT 180
 OY 181 GTATCCGTACCGACGATGAGAGCGGCGAGAAAGCTTCAAACTCATCAACAGAGGTTAAC 240
 Db 181 GTATCCGTACCGACGATGAGAGCGGCGAGAAAGCTTCAAACTCATCAACAGAGGTTAAC 240
 OY 241 GTGCCGCTGTGGCTGACATTCACATTCGATTCGATTCGCTGAAAGTACGGGAATAC 300
 Db 241 GTGCCGCTGTGGCTGACATTCACATTCGATTCGATTCGCTGAAAGTACGGGAATAC 300
 OY 301 GGGCGATGTTCTGCGTAACTTAACTTACCGTATTCGATTAAGAGAGCGTATTCGATG 360
 Db 301 GGGCGATGTTCTGCGTAACTTAACTTACCGTATTCGATTAAGAGAGCGTATTCGATG 360
 OY 361 GTGGTTGACTGTGCGCGGATAAAAAATTCGATTCGATTCGCTTAAAGCGCGGATCG 420
 Db 361 GTGGTTGACTGTGCGCGGATAAAAAATTCGATTCGATTCGCTTAAAGCGCGGATCG 420
 OY 421 CTGGAAGAAAGATCTCAAGAAAGTATGGGACCGGCGGCGGCTTCTGGAATCT 480
 Db 421 CTGGAAGAAAGATCTCAAGAAAGTATGGGACCGGCGGCGGCTTCTGGAATCT 480
 OY 481 GGCATGCTCATGTTGATGATCTGATGCGCTGAATCTCATGCTCAAAAGTACAGCGTG 540
 Db 481 GGCATGCTCATGTTGATGATCTGATGCGCTGAATCTCATGCTCAAAAGTACAGCGTG 540
 OY 541 AAAGCTGTGACGTTCTTCTGCGCTGTTGAGTCTTATGCTTGGCGCAAAAGATGAT 600
 Db 541 AAAGCTGTGACGTTCTTCTGCGCTGTTGAGTCTTATGCTTGGCGCAAAAGATGAT 600
 OY 601 CAGCGTTGCTGCTGGGAGAGACCGGAGCGGCTGCGGACAGCGGCGAGTAAATCC 660
 Db 601 CAGCGTTGCTGCTGGGAGAGACCGGAGCGGCTGCGGACAGCGGCGAGTAAATCC 660
 OY 661 GCGATTGGTTAGGTCTGCGCTGTGTGAAGGATCGGCGACAGCGTGGCGGATGCTG 720
 Db 661 GCGATTGGTTAGGTCTGCGCTGTGTGAAGGATCGGCGACAGCGTGGCGGATGCTG 720
 OY 721 GCGGCGATCCGCTGCAAGAGATCAAAAGTCTTGAATTTTGAATCGCTGCTATC 780
 Db 721 GCGGCGATCCGCTGCAAGAGATCAAAAGTCTTGAATTTTGAATCGCTGCTATC 780
 OY 781 CGTTCCGAGAGGATCAACTGATCGCTCCGACCTGTTCCGTCAGAGAAATTTATGTT 840
 Db 781 CGTTCCGAGAGGATCAACTGATCGCTCCGACCTGTTCCGTCAGAGAAATTTATGTT 840
 OY 841 ATGGGTAAGGTTAAGCGGCTGAGCAAGCGCTGGAAGATATCATCACTCCGATGAGCTT 900
 Db 841 ATGGGTAAGGTTAAGCGGCTGAGCAAGCGCTGGAAGATATCATCACTCCGATGAGCTT 900
 OY 901 TCGATTATGGGCTGGGTGATGGCCAGAGTGAAGCGCTGGTTTCAACCTCGCGCTC 960
 Db 901 TCGATTATGGGCTGGGTGATGGCCAGAGTGAAGCGCTGGTTTCAACCTCGCGCTC 960
 OY 961 ACCGGCGGCAACAGAAAAAGGGCTCTATGAAGATGGCGGCAAGAGCGCTGTGAC 1020
 Db 961 ACCGGCGGCAACAGAAAAAGGGCTCTATGAAGATGGCGGCAAGAGCGCTGTGAC 1020

Db 961 ACCGGCGGCAACAGAAAAAGGGCTCTATGAAGATGGCGGCAAGAGCGCTGTGAC 1020
 OY 1021 AACACGATATGATGACAGCAGCTGGAAGCAGCAGTGTGGGAAGGACGACTGAC 1080
 Db 1021 AACACGATATGATGACAGCAGCTGGAAGCAGCAGTGTGGGAAGGACGACTGAC 1080
 OY 1081 GAAGCGCTGCAATTTGACGTTACAGAGTTGAAAAATTA 1119
 Db 1081 GAAGCGCTGCAATTTGACGTTACAGAGTTGAAAAATTA 1119

RESULT 2
 AX038912 1119 bp DNA linear PAT 16-NOV-2000
 LOCUS AX038912
 DEFINITION Sequence 29 from Patent WO0061793.
 ACCESSION AX038912
 VERSION AX038912.1 GI:11228221
 KEYWORDS
 SOURCE
 ORGANISM
 Escherichia coli.
 Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

REFERENCE
 1 (bases 1 to 1119)
 AUTHORS Loferer, H. and Jacob, A.
 TITLE Novel method for identifying antibacterial compounds
 JOURNAL Patent: WO 0061793-A 29-19-OCT-2000;
 GPC BIOTECH AG (DE); LOFERER HANNES (DE); JACOBI ALEXANDER (DE)

FEATURES
 source
 1..1119
 /organism="Escherichia coli"
 /db_xref="taxon:562"

BASE COUNT 272 a 279 c 317 g 251 t
 ORIGIN

Query Match 100.0%; Score 1119; DB 6; Length 1119;
 Best Local Similarity 100.0%; Pred. No. 9.9e-271;
 Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGCATACACAGGCTCATATCAAGCTAGAAAAATCAACAGCTATTTACGTTGGGAATGTG 60
 Db 1 ATGCATACACAGGCTCATATCAAGCTAGAAAAATCAACAGCTATTTACGTTGGGAATGTG 60
 OY 61 CCGATTGGCGATGTGCTGCATCGCCGTACAGTCCATACCAATACGGGTACAGACAGAC 120
 Db 61 CCGATTGGCGATGTGCTGCATCGCCGTACAGTCCATACCAATACGGGTACAGACAGAC 120
 OY 121 GTGGAAGCAAGGCTCAATGAATCAAGGGCGTGAAGCGGTTGGCGCTGATATCGCTT 180
 Db 121 GTGGAAGCAAGGCTCAATGAATCAAGGGCGTGAAGCGGTTGGCGCTGATATCGCTT 180
 OY 181 GTATCCGTACCGACGATGAGAGCGGCGAGAAAGCTTCAAACTCATCAACAGAGGTTAAC 240
 Db 181 GTATCCGTACCGACGATGAGAGCGGCGAGAAAGCTTCAAACTCATCAACAGAGGTTAAC 240
 OY 241 GTGCCGCTGTGGCTGACATTCACATTCGATTCGATTCGCTGAAAGTACGGGAATAC 300
 Db 241 GTGCCGCTGTGGCTGACATTCACATTCGATTCGATTCGCTGAAAGTACGGGAATAC 300
 OY 301 GGGCGATGTTCTGCGTAACTTAACTTACCGTATTCGATTAAGAGAGCGTATTCGATG 360
 Db 301 GGGCGATGTTCTGCGTAACTTAACTTACCGTATTCGATTAAGAGAGCGTATTCGATG 360
 OY 361 GTGGTTGACTGTGCGCGGATAAAAAATTCGATTCGATTCGCTTAAAGCGCGGATCG 420
 Db 361 GTGGTTGACTGTGCGCGGATAAAAAATTCGATTCGATTCGCTTAAAGCGCGGATCG 420
 OY 421 CTGGAAGAAAGATCTCAAGAAAGTATGGGACCGGCGGCGGCTTCTGGAATCT 480
 Db 421 CTGGAAGAAAGATCTCAAGAAAGTATGGGACCGGCGGCGGCTTCTGGAATCT 480
 OY 481 GGCATGCTCATGTTGATGATCTGATGCGCTGAATCTCATGCTCAAAAGTACAGCGTG 540
 Db 481 GGCATGCTCATGTTGATGATCTGATGCGCTGAATCTCATGCTCAAAAGTACAGCGTG 540

QY 541 AAAGCGTCTGACGCTCTCTCCGCGTGGTGAAGCTTATGTTTGGTGGCAAAACAGATCAT 600
 DB 541 AAAGCGTCTGACGCTCTCTCCGCGTGGTGAAGCTTATGTTTGGTGGCAAAACAGATCAT 600
 QY 601 CAGCGGTTGATCTGGGGATCACCGAAGCCGCTGGCGCGCAGCGGGGCGAGTAAATCC 660
 DB 601 CAGCGGTTGATCTGGGGATCACCGAAGCCGCTGGCGCGCAGCGGGGCGAGTAAATCC 660
 QY 661 GCCATTGTTAGTCTGCTGCTCTCTGAAGGATCGGCGACACGCTGGCGGATTCGCTG 720
 DB 661 GCCATTGTTAGTCTGCTGCTCTCTGAAGGATCGGCGACACGCTGGCGGATTCGCTG 720
 QY 721 GCGGCGCATCCGCTCGAAGAGATCAAGTCCGTTTCATATTTGAAATGCTGCGATC 780
 DB 721 GCGGCGCATCCGCTCGAAGAGATCAAGTCCGTTTCATATTTGAAATGCTGCGATC 780
 QY 781 CGTTGCGAGGATCACTTCATCGCTGCGCGACCTGTTGCGCTCAAGGATTTGATGTT 840
 DB 781 CGTTGCGAGGATCACTTCATCGCTGCGCGACCTGTTGCGCTCAAGGATTTGATGTT 840
 QY 841 ATCGGTACGTTAAACGCGCTGGAGCAACGCTGGAAGATATCATCTCCGATGGAGCTT 900
 DB 841 ATCGGTACGTTAAACGCGCTGGAGCAACGCTGGAAGATATCATCTCCGATGGAGCTT 900
 QY 901 TCGATTATCGGCTGCGTGAATGAGCCAGTGAAGCGCGTGTTCATCACTGCGCGTC 960
 DB 901 TCGATTATCGGCTGCGTGAATGAGCCAGTGAAGCGCGTGTTCATCACTGCGCGTC 960
 QY 961 ACCGCGGCAACAGAAAAAGCGGCTCTATGAGATGGCGTGGCGAAGACCGTCTGGAC 1020
 DB 961 ACCGCGGCAACAGAAAAAGCGGCTCTATGAGATGGCGTGGCGAAGACCGTCTGGAC 1020
 QY 1021 AACCAAGATATGATGACCACTGGAAGCAGCATTCCTGGGAAAGCAGTCACTGGAC 1080
 DB 1021 AACCAAGATATGATGACCACTGGAAGCAGCATTCCTGGGAAAGCAGTCACTGGAC 1080
 QY 1081 GAAGCGCTCGAATGAGCTTCAGCAGGTTGAAAAATAA 1119
 DB 1081 GAAGCGCTCGAATGAGCTTCAGCAGGTTGAAAAATAA 1119
 RESULT 3
 AX050487 1119 bp DNA linear PAT 12-JAN-2001
 LOCUS Sequence 1 from Patent WO0072022.
 ACCESSION AX050487
 VERSION AX050487.1 GI:12226691
 KEYWORDS
 SOURCE
 ORGANISM
 Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 1 (bases 1 to 1119)
 JOURNAL
 Jomaa, H.
 Use of genes of the deoxy-d-xylose phosphate biosynthetic pathway
 for altering the concentration of isoprenoid
 Patent: WO 0072022-A 1 30-NOV-2000;
 Jomaa, Hassan (DE)
 FEATURES
 source
 1..1119
 Location/Qualifiers
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 1..1119
 /note="unnamed protein product"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC21801.1"
 /db_xref="GI:12226692"
 /translation="MHNOAPIORRSTRIYGNVPISGDPAIVASMTNRTTVEAT
 VNOTKALEKRGADIVRVSPPTMDAAEAKLIKQOVNVPVLEDDHDEYIAIKVAEYV
 DCLRTNPNIGNEKIRNVPVCAKRNKPIPIKIGVAGSLVEDLDEKTEPPQALLES
 AMRHVDHLDRLNFDQFKVSKASDVFLAVESTRLAKOIIDPLHLGITTEAGANSNAV

KSAIGLLISBGIIDTLRVSIAADPEEIKVGFLLKSLRISRGINFIACTCSRO
 EPDVIETVALAEQREDIITPMDSVSIIGVNPBGLALVSTLGTGNKRSGLYEDGV
 RKRLNNMIMDLEARIKAKSOLDEARIDVQVEK"
 BASE COUNT 272 a 279 c 317 g 251 t
 ORIGIN
 Query Match 100.0%; Score 1119; DB 6; Length 1119;
 Best Local Similarity 100.0%; Pred. No. 9,9e-271;
 Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGCATACCAAGGCTCCCAATTCAGAGTGAAGATCAACAGCATATTTACGTTGGAAATCTG 60
 DB 1 ATGCATACCAAGGCTCCCAATTCAGAGTGAAGATCAACAGCATATTTACGTTGGAAATCTG 60
 QY 61 CCGATTGGCGATGGTGGTCCCATGCGGTACAGTCACATGACCATPAGCGGTAGCAGAC 120
 DB 61 CCGATTGGCGATGGTGGTCCCATGCGGTAGTCACATGACCATPAGCGGTAGCAGAC 120
 QY 121 GTGCAAGCAACGCTCAATCAATCAATCAAGCGCTGGAAGCGCTTGGCGCTGATATCGTCGT 180
 DB 121 GTGCAAGCAACGCTCAATCAATCAATCAAGCGCTGGAAGCGCTTGGCGCTGATATCGTCGT 180
 QY 181 GTATCCGTACCGACAGATGAGCGCGGAGAGCGTTCAACATCTCAACAGCAGGTTATAC 240
 DB 181 GTATCCGTACCGACAGATGAGCGCGGAGAGCGTTCAACATCTCAACAGCAGGTTATAC 240
 QY 241 GTGCGCGGTGGTGGTGCATCCATCTGACATGACATTCGATTCGCTGAAAGTATGCGGAATAC 300
 DB 241 GTGCGCGGTGGTGGTGCATCCATCTGACATGACATTCGATTCGCTGAAAGTATGCGGAATAC 300
 QY 301 GCGCTCGATTGCTGCTGATTAATACCTGGCAATATCGGTAATGAAGACGTTATTCGATG 360
 DB 301 GCGCTCGATTGCTGCTGATTAATACCTGGCAATATCGGTAATGAAGACGTTATTCGATG 360
 QY 361 GTGTTGACTGTGGCGGAGTAAAGCAATTCGATTCGCTGATTTGGCGTTAATGCGCGATCG 420
 DB 361 GTGTTGACTGTGGCGGAGTAAAGCAATTCGATTCGCTGATTTGGCGTTAATGCGCGATCG 420
 QY 421 CTGCAAAAGATCTGCAAGAAAGTATGGCGAAGCGAGCGCGGCTTCTGGAATCT 480
 DB 421 CTGCAAAAGATCTGCAAGAAAGTATGGCGAAGCGAGCGCGGCTTCTGGAATCT 480
 QY 481 GCCATGCTCATGTTGATCATCTGATCGCCTGAACCTTCATCACTTAAAGTCAAGCTG 540
 DB 481 GCCATGCTCATGTTGATCATCTGATCGCCTGAACCTTCATCACTTAAAGTCAAGCTG 540
 QY 541 AAAGCGCTGACGCTCTCTCGCTGCTGATGATCTTATGCTTTGCTGGCAAAACAGATCAT 600
 DB 541 AAAGCGCTGACGCTCTCTCGCTGCTGATGATCTTATGCTTTGCTGGCAAAACAGATCAT 600
 QY 601 CAGCGGTTGATCTGGGGATCACCGAAGCCGCTGGCGCGCAGCGGGGCGAGTAAATCC 660
 DB 601 CAGCGGTTGATCTGGGGATCACCGAAGCCGCTGGCGCGCAGCGGGGCGAGTAAATCC 660
 QY 661 GCCATTGTTAGTCTGCTGCTCTCTGAAGGATCGGCGACACGCTGGCGGATTCGCTG 720
 DB 661 GCCATTGTTAGTCTGCTGCTCTCTGAAGGATCGGCGACACGCTGGCGGATTCGCTG 720
 QY 721 GCGGCGCATCCGCTCGAAGAGATCAAGTCCGTTTCATATTTGAAATGCTGCGATC 780
 DB 721 GCGGCGCATCCGCTCGAAGAGATCAAGTCCGTTTCATATTTGAAATGCTGCGATC 780
 QY 781 CGTTGCGAGGATCACTTCATCGCTGCGCGACCTGTTGCGCTCAAGGATTTGATGTT 840
 DB 781 CGTTGCGAGGATCACTTCATCGCTGCGCGACCTGTTGCGCTCAAGGATTTGATGTT 840
 QY 841 ATCGGTACGTTAAACGCGCTGGAGCAACGCTGGAAGATATCATCTCCGATGGAGCTT 900
 DB 841 ATCGGTACGTTAAACGCGCTGGAGCAACGCTGGAAGATATCATCTCCGATGGAGCTT 900
 QY 901 TCGATTATCGGCTGCGTGAATGAGCCAGTGAAGCGCGTGTTCATCACTGCGCGTC 960
 DB 901 TCGATTATCGGCTGCGTGAATGAGCCAGTGAAGCGCGTGTTCATCACTGCGCGTC 960

OY 961 ACCGGCGCAACAGAAAAGCGCCTCTATGAGATGGCGTGGCAAGACCGTCTGGAC 1020
 Db 961 ACCGGCGCAACAGAAAAGCGCCTCTATGAGATGGCGTGGCAAGACCGTCTGGAC 1020
 OY 1021 AACACGATATGATGACGCGCTGGAAGCAGCATTCGTGGAAAGCAGTCTGGAC 1080
 Db 1021 AACACGATATGATGACGCGCTGGAAGCAGCATTCGTGGAAAGCAGTCTGGAC 1080
 OY 1081 GAAGCGCGTGAATTCAGCTGACAGGTTGAAATAA 1119
 Db 1081 GAAGCGCGTGAATTCAGCTGACAGGTTGAAATAA 1119
 RESULT 4
 AX393943 1119 bp DNA linear PAT 23-MAR-2002
 LOCUS Sequence 3 from Patent WO0212478.
 DEFINITION AX393943
 ACCESSION AX393943
 VERSION AX393943.1 GI:19701905
 KEYWORDS
 SOURCE Escherichia coli.
 ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
 REFERENCE
 AUTHORS Boronac, A., Campos, N., Rodriguez-Concepcion, M., Rohmer, M., Seeman, M., Valentin, H.E., Venkatesh, T.V. and Venkatesh, M.
 TITLE Methyl-d-erythritol phosphate pathway genes
 JOURNAL Patent: WO 0212478-A 3 14-FEB-2002; Monsanto Technology LLC (US)
 FEATURES
 source Location/Qualifiers
 1. 1119
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 1. 1119
 /note="unnamed protein product"
 /codon_start=1
 /transl_table=11
 /protein_id="CAD29065.1"
 /db_xref="gi:19701906"
 /translation="MHNOAPIORRSTRIVGNVPIGDAPITAVOSMTNRTTDEAT
 VNOIKALEVAGADIVREVSPTDAAEAFLIKOOVNPVIAIDHPEYRAIVAEVY
 DCLRNGTGNERTRMVYVDCARDKNIPIRIGVNAAGSEKIDOEYGGPTDALLS
 AMRHVDHEDLNDFKVSADVFLAESTRLAKQIDPILHIGITAGARSAGV
 KSLIGLIELSEIGIDTLRSLAADVEIKRFLDLKRLRSRGINIACPTCSRO
 EEDVIGTVALDREDDITPMDVSIIGCVNGPGLAIVSTLGTGNGKSSLYEDGV
 RDRIDNNMDIOLFIARIRAKASOLDEARIDVQVEK"
 BASE COUNT 272 a 279 c 317 g 251 t
 ORIGIN
 Query Match 100.0%; Score 1119; DB 6; Length 1119;
 Best Local Similarity 100.0%; Pred. No. 9,9e-271;
 Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 241 GTGCCGCTGGGCTGACATTCACCTTCGACTATGCAATTGGCTGAAATAGCGGAATAC 300
 OY 301 GGCGTCGATGTCGCGATTAACCCGCAATTCGGTATGAAAGCGCATTCGATG 360
 Db 301 GGCGTCGATGTCGCGATTAACCCGCAATTCGGTATGAAAGCGCATTCGATG 360
 OY 361 GTGGTACTGTGGCGCGATTAACCATTCGATTCGATTCGATTCGATTCGATTCG 420
 Db 361 GTGGTACTGTGGCGCGATTAACCATTCGATTCGATTCGATTCGATTCGATTCG 420
 OY 421 CTGGAANAAGATTCGCAAGAAAGATATGGGAACCGGCGGAGGCTTCGAAATC 480
 Db 421 CTGGAANAAGATTCGCAAGAAAGATATGGGAACCGGCGGAGGCTTCGAAATC 480
 OY 481 GCCATGCGTATGTCGATTCATTCGATTCGATTCGATTCGATTCGATTCGATTCG 540
 Db 481 GCCATGCGTATGTCGATTCATTCGATTCGATTCGATTCGATTCGATTCGATTCG 540
 OY 541 AAAGCGTCTGACGTCCTTCGCTGCTGTTGAATTCGATTCGATTCGATTCGATTCG 600
 Db 541 AAAGCGTCTGACGTCCTTCGCTGCTGTTGAATTCGATTCGATTCGATTCGATTCG 600
 OY 601 CAGCGCTGATCTCGGGGATCCAGGAACCGCGTGGTGGCGGAGCGGCGATTAATCC 660
 Db 601 CAGCGCTGATCTCGGGGATCCAGGAACCGCGTGGTGGCGGAGCGGCGATTAATCC 660
 OY 661 GCCATTGCTTAAAGTCTGCTGCTGCTGCTGAAGGATCGGCGGCGGCGGATTCGATTCG 720
 Db 661 GCCATTGCTTAAAGTCTGCTGCTGCTGCTGAAGGATCGGCGGCGGCGGATTCGATTCG 720
 OY 721 GCGGCGGATCCGCTCGAAGAGATCAAAAGTGGTTCGATTCGATTCGATTCGATTCG 780
 Db 721 GCGGCGGATCCGCTCGAAGAGATCAAAAGTGGTTCGATTCGATTCGATTCGATTCG 780
 OY 781 CGTTGCGAGGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 840
 Db 781 CGTTGCGAGGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 840
 OY 841 ATGCGTACGCTTAACGCGCTGAGCAAGCGCTGGAAGATATCACTCCGATGACGTT 900
 Db 841 ATGCGTACGCTTAACGCGCTGAGCAAGCGCTGGAAGATATCACTCCGATGACGTT 900
 OY 901 TCATTATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
 Db 901 TCATTATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
 OY 961 ACCGGCGGCAACAGAAAGCGGCTCTATGAAGATGGGCTGGCAAGACCGTCTGGAC 1020
 Db 961 ACCGGCGGCAACAGAAAGCGGCTCTATGAAGATGGGCTGGCAAGACCGTCTGGAC 1020
 OY 1021 AACACGATATGATTCGACGAGCTGGAAGCAGCATTCGTGGAAAGCAGTCTGGAC 1080
 Db 1021 AACACGATATGATTCGACGAGCTGGAAGCAGCATTCGTGGAAAGCAGTCTGGAC 1080
 OY 1081 GAAGCGCGTGAATTCAGCTTACGCAAGTTGAAATAA 1119
 Db 1081 GAAGCGCGTGAATTCAGCTTACGCAAGTTGAAATAA 1119
 RESULT 5
 ECSCP ECGPE 1697 bp DNA linear BCT 08-DEC-1993
 LOCUS E.coli gcpE gene.
 DEFINITION X64451.1 S43432
 ACCESSION X64451.1 GI:11540
 VERSION gcpE gene.
 KEYWORDS gcpE gene.
 SOURCE Escherichia coli.
 ORGANISM Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
 REFERENCE 1 (bases 1 to 1697)
 AUTHORS Parker, J.
 TITLE Direct Submission

JOURNAL Submitted (17-FEB-1992) J. Parker, Southern Illinois Univ, Dept of Microbiology, S I U, Carbondale IL 62901, USA

REFERENCE 2 (bases 1 to 1697)

AUTHORS Baker, J., Franklin, D. B. and Parker, J.

TITLE Sequence and characterization of the *gcpE* gene of *Escherichia coli*

JOURNAL FEMS Microbiol. Lett. 94, 175-180 (1992)

COMMENT See also J01629 & M11843.

FEATURES

Location/Qualifiers

1..1697

/organism="Escherichia coli"

/strain="K12"

/db_xref="taxon:562"

/sub_clone="PSE411"

/cell_line="JF298"

/clone_1lb="Lambda transducing phage dguab-2"

<1..508

/note="ORF"

/codon_start=2

/transl_table=1

/protein_id="CAA45782.1"

/db_xref="GI:41541"

/db_xref="SWISS-PROT:P27434"

/translation="TTTDPATSTPPASVDTATNTOTPAVTAAPAVDPOQNAVVP.
/transl_table=1
SQANDVATATPAATTPDGAAPLPDQGVTPPADPVALVNFATDCMLEVDTAT
GKLFSGMQRKDNLTGAPVYKLTIGAPNAVYIQGKRPVDSRTIRTNQVARI
NAEOSPAQ"

535..1653

/gene="gcpE"

535..1653

/gene="gcpE"

/codon_start=1

/transl_table=1

/protein_id="CAA45783.1"

/db_xref="GI:41542"

/db_xref="SWISS-PROT:P27433"

/translation="MHNQPIORRSTRIYGNVPIGDGAPVAVOSMTNRTTDEAT
/transl_table=1
VNOIKLEAVGADIVRYSVPTDAAEFKLQOVNPLVDIHFEDVIAIKVAYGY
DCIRINPENGEBERIRMYPCARDKNTPIRIGVAGSLEKDLQKYEPPQALLES
AMRHVDHLRLSDQFVSKASDVPLAVESRYLLAKIIDPILHIGTIEAGASGAV
KSAIIGDLISBGIPTLVSLIADVBEIKYGPDLIKSLIRSRGINFIACPCSRQ
EFVIGTVNALEQREDIITPMDVSLIGCVVGPBALVSTLIGVIGKSKGLVEDGY
RKDLDNMIDOLEARIRAKASOLDEARIDVOVER"

CDS

gene

535..1653

535..1653

CDS

BASE COUNT 417 a 446 c 469 g 365 t

ORIGIN

Query Match 100.0%; Score 1119; DB 1; Length 1697;

Best Local Similarity 100.0%; Pred. No. 1e-270; Indels 0; Gaps 0;

Matches 1119; Conservative 0; Mismatches 0;

OY 1 ATGCTAACCAAGGCTCCCAATTCACGTAGAAATCAACAGCTATTACGTTGGGATGTG 60

DB 535 ATGCTAACCAAGGCTCCCAATTCACGTAGAAATCAACAGCTATTACGTTGGGATGTG 594

OY 61 CCGATTGGGATGTGCTCCATGCGCGGTACAGTCATGACCAATACGCGGTACAGACAG 120

DB 595 CCGATTGGGATGTGCTCCATGCGCGGTACAGTCATGACCAATACGCGGTACAGACAG 654

OY 121 GTTCGAAGCAAGCTCAATCAATCAAGGCGGTGAAGCGGTGGCGGTATGTCGCGT 180

DB 655 GTTCGAAGCAAGCTCAATCAATCAAGGCGGTGAAGCGGTGGCGGTATGTCGCGT 714

OY 181 GTATCCGTACCGACGATGACGCGGAGAAAGCGTTCAAACTCATCAACAGAGGTAAAC 240

DB 715 GTATCCGTACCGACGATGACGCGGAGAAAGCGTTCAAACTCATCAACAGAGGTAAAC 774

OY 241 GTCCCGCTGTGCTGACATCCACTTCGACTATCGCATTCGCGCTGAAGTACGGAATAC 300

DB 775 GTCCCGCTGTGCTGACATCCACTTCGACTATCGCATTCGCGCTGAAGTACGGAATAC 834

OY 301 GGGGTGATGTTGCTGCTATTAACCTGCGAATATGCTGATGAAGAGCGTATTGGCATG 360

DB 835 GGGGTGATGTTGCTGCTATTAACCTGCGAATATGCTGATGAAGAGCGTATTGGCATG 894

OY 361 GTGCTGACTGTGGCGCGATTAATAAACATTCGATCCGATATTGGCGTTAACCGCGGATG 420

DB 895 GTGCTGACTGTGGCGCGATTAATAAACATTCGATCCGATATTGGCGTTAACCGCGGATG 954

OY 421 CTGGAAAAAGATCTGCAAGAAAAATATGCGAACCGACGCGCGAGCGTGTGATCTG 480

DB 955 CTGGAAAAAGATCTGCAAGAAAAATATGCGAACCGACGCGCGAGCGTGTGATCTG 1014

OY 481 GCCATGCTCATGTTGATCATCTGATCCGCTGAACCTTCGATCAGTTCAAAAGTCAGGCTG 540

DB 1015 GCCATGCTCATGTTGATCATCTGATCCGCTGAACCTTCGATCAGTTCAAAAGTCAGGCTG 1074

OY 541 AAAGCTGTGACGTCTTCTCTGCTGCTGATGATGCTTATGCTTTGCTGGCAAAAAGATCAT 600

DB 1075 AAAGCTGTGACGTCTTCTCTGCTGCTGATGATGCTTATGCTTTGCTGGCAAAAAGATCAT 1134

OY 601 CAGCGCTGATCTGCGGATCAACGAGCGGTGGTGGCGCGACGCGGCGGTAAATTC 660

DB 1135 CAGCGCTGATCTGCGGATCAACGAGCGGTGGTGGCGCGACGCGGCGGTAAATTC 1194

OY 661 GCGATGTTAGTCTGCTGCTGCTGCTGAAGCATGCGCGACAGCTGCGGTATCGCTG 720

DB 1195 GCGATGTTAGTCTGCTGCTGCTGCTGAAGCATGCGCGACAGCTGCGGTATCGCTG 1254

OY 721 GCGCGCATTCGCTGCAAGAGATCAAAAGCGGTTTCGATATTGAAATCGCTGCTATC 780

DB 1255 GCGCGCATTCGCTGCAAGAGATCAAAAGCGGTTTCGATATTGAAATCGCTGCTATC 1314

OY 781 CGTTGCGAGGATCAACTTCATCGCTGCCGACCTGTTGCGCTGAGAAATTTATGTT 840

DB 1315 CGTTGCGAGGATCAACTTCATCGCTGCCGACCTGTTGCGCTGAGAAATTTATGTT 1374

OY 841 ATCGGTACGCTTAAACGCGCTGAGCAACCGCTGGAAGATATCATCTCGATGACGTT 900

DB 1375 ATCGGTACGCTTAAACGCGCTGAGCAACCGCTGGAAGATATCATCTCGATGACGTT 1434

OY 901 TCGATTTCGCGCTGCGGTGTAATGCGCGAGTGAAGCGCGCTGTTCTACACTGCGCTG 960

DB 1435 TCGATTTCGCGCTGCGGTGTAATGCGCGAGTGAAGCGCGCTGTTCTACACTGCGCTG 1494

OY 961 ACCGCGGCAACAGAAAAAGCGGCTTATGAAGATGCGGTGGCAAAAGACGCTGAGAC 1020

DB 1495 ACCGCGGCAACAGAAAAAGCGGCTTATGAAGATGCGGTGGCAAAAGACGCTGAGAC 1554

OY 1021 AACAAGATATGATGACACGACGCTGAAGCAAGCATTCGTGCGAAAGCAGCTGAGTGAC 1080

DB 1555 AACAAGATATGATGACACGACGCTGAAGCAAGCATTCGTGCGAAAGCAGCTGAGTGAC 1614

OY 1081 GAAGCGCTGCAATGACGTTACAGAGTTGAAATTA 1119

DB 1615 GAAGCGCTGCAATGACGTTACAGAGTTGAAATTA 1653

RESULT 6

AE000338/c 13176 bp DNA linear BCT 01-DEC-2000

LOCUS Escherichia coli K12 MG1655 section 228 of 400 of the complete genome.

DEFINITION AE000338.1 GI:1788862

ACCESSION AE000338.1

VERSION AE000338.1

KEYWORDS

SOURCE

ORGANISM

Escherichia coli K12.

Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

REFERENCE 1 (bases 1 to 13176)

AUTHORS Blattner, F.R., Plunkett, G. III, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. and Shao, Y.

TITLE The complete genome sequence of *Escherichia coli* K-12

JOURNAL Science 277 (5331), 1453-1474 (1997)

MEDLINE 97426617


```

/db_xref="taxon:562"
1..1119
/gene="ispG"
1..1119
/gene="ispG"
/function="converts 2C-methyl-D-erythritol
2,4-cyclodiphosphate into 1-hydroxy-2-methyl-2-(E)-butenyl
4-diphosphate, last but one step of the
mevalonate-independent pathway for isoprenoid
biosynthesis"
/note="ispG"
/codon_start=1
/transl_table=11
/product="1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate
synthase"
/protein_id="AAK53460.1"
/db_xref="GI:17978528"
/translation="MHNQAPLQRRKSTRIVGNPLIGDGAPIAVOSMTNTRTDVEAT
VNOIKALEYAGADIVRSVPTMADAEAFKIKOYVNPPLVADHFDRIALKAIEYV
DCLRINFTGNERNIRMYVDCARDKNIPRIQVNASLEKDLQEKYGEPTPOLLES
AMRHVDHEDLNEDQFVSVKASDVLAESYRLAKQIDQPHLGLTEAGARGAV
KSAIGLEELSEGIDTLRVSLADPVEIKVGDILKSLRISRGINFACPCSHQ
EFVDIGTVNLEORLEDITPMDSIIGCVNNGEALVSTLGVTVGNKKSGILEDSQ
RKRLDNDMDIDLEARIKAKASOLDEARRIDVQOVER"
BASE COUNT      272 a      280 c      315 g      252 t
ORIGIN

```

```

Query Match      97.7% Score 1093.4; DB 1; Length 1119;
Best Local Similarity 98.6%; Pred. No. 2,7e-264;
Matches 1103; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```

```

QY 1 ATGCATACACAGGCTCCAAATCAAGTAGAATAATCAACAGTATTACGTGGGAATGTG 60
DB 1 ATGCATACACAGGCTCCAAATCAAGTAGAATAATCAACAGTATTACGTGGGAATGTG 60
QY 61 CCGATTGGCGATGCTGCGATCGCCGTACAGTCATGACCAATAGCCGTACAGAC 120
DB 61 CCGATTGGCGATGCTGCGATCGCCGTACAGTCATGACCAATAGCCGTACAGAC 120
QY 121 GTGCAAGCAAGCGTCATGAAATCAAGCGGTGGAAGCGCTGGCGCTGATATGTCCTG 180
DB 121 GTGCAAGCAAGCGTCATGAAATCAAGCGGTGGAAGCGCTGGCGCTGATATGTCCTG 180
QY 181 GTATCCGTACCGAGATGAGCGCGGCAAGACGTTCAACATCAACAGCAGCATTAAC 240
DB 181 GTATCCGTACCGAGATGAGCGCGGCAAGACGTTCAACATCAACAGCAGCATTAAC 240
QY 241 GTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
DB 241 GTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
QY 301 GGGCGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 301 GGGCGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 361 GTGGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 361 GTGGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 421 CTGCAAAAACATCTGCAAGAAAGATATGCGAAGCAGCGCGGAGCGGTTCTGGAATCT 480
DB 421 CTGCAAAAACATCTGCAAGAAAGATATGCGAAGCAGCGCGGAGCGGTTCTGGAATCT 480
QY 481 GCCATGCGCATGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCA 540
DB 481 GCCATGCGCATGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCA 540
QY 541 AAAGCGTCAAGCTCTTCCGCTGTTGAGTTATGTTGCTGCAAAAACAGATGAT 600
DB 541 AAAGCGTCAAGCTCTTCCGCTGTTGAGTTATGTTGCTGCAAAAACAGATGAT 600
QY 601 CAGCGCTGATCTGGGAGTACCGAAGCGGCTGCTGCGGCAAGCGGCGCAGTAAATCC 660
DB 601 CAGCGCTGATCTGGGAGTACCGAAGCGGCTGCTGCGGCAAGCGGCGCAGTAAATCC 660

```

```

DB 601 CAGCGCTGATCTGGGAGTACCGAAGCGGCTGCTGCGGCAAGCGGCGCAGTAAATCC 660
QY 661 GCCATTGGTTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 661 GCCATTGGTTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 721 GCGGCGCATCCGCGTCAAGAGATCAAAAGTCCGTTGCAATTTGAAATGCTTCGATC 780
DB 721 GCGGCGCATCCGCGTCAAGAGATCAAAAGTCCGTTGCAATTTGAAATGCTTCGATC 780
QY 781 CGTTGCGAGGATCACTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB 781 CGTTGCGAGGATCACTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 841 ATCGGTACAGTTAAGCGGCTGAGCAAGCGGCTGAGCAAGTATCATCATCTCCGATGAG 900
DB 841 ATCGGTACAGTTAAGCGGCTGAGCAAGCGGCTGAGCAAGTATCATCATCTCCGATGAG 900
QY 901 TCGATTATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
DB 901 TCGATTATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 961 ACCGGCGGCAACAGAAAGAGCGGCTCTATGAGATGCGCTGCGCAAAAGCCGTGAGAC 1020
DB 961 ACCGGCGGCAACAGAAAGAGCGGCTCTATGAGATGCGCTGCGCAAAAGCCGTGAGAC 1020
QY 1021 AACACAGATATGATGACGACAGCTGAGCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTG 1080
DB 1021 AACACAGATATGATGACGACAGCTGAGCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTG 1080
QY 1081 GAAGCGGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1119
DB 1081 GAAGCGGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1119

```

RESULT 8
 AE005481/c
 LOCUS
 DEFINITION
 Escherichia coli O157:H7 EDL933 genome, contig 3 of 3, section 100
 of 290.
 ACCESSION
 AE005481 AE005174
 VERSION
 AE005481.1 GI:12516908
 KEYWORDS
 SOURCE
 ORGANISM
 Escherichia coli O157:H7 EDL933.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

REFERENCE
 AUTHORS
 1 (bases 1 to 11521)
 Perna N.T., Plunkett, G. III, Burland, V., Mau, B., Glasner, J.D., Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A., Postel, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grobeck, E.J., Davis, N.W., Lim, A., Dimalanta, E., Potamousis, K., Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C., Welch, R.A. and Blattner, F.R.
 TITLE
 Genome sequence of enterohemorrhagic Escherichia coli O157:H7
 JOURNAL
 Nature 409 (6619), 529-533 (2001)
 MEDLINE
 21074935
 PUBMED
 11206551

REFERENCE
 AUTHORS
 2 (bases 1 to 11521)
 Perna, N.T., Plunkett, G. III, Burland, V., Mau, B., Glasner, J.D., Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A., Postel, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grobeck, E.J., Davis, N.W., Lim, A., Dimalanta, E., Potamousis, K., Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C., Welch, R.A. and Blattner, F.R.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (22-OCT-2000) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
 FEATURES
 source
 1..11521
 location/Qualifiers
 /organism="Escherichia coli O157:H7 EDL933"
 /strain="EDL933"
 /serotype="O157:H7"

```

gene
/db_xref="taxon:155864"
/note="enterohemorrhagic"
/complement(86..1204)
/gene="gcpE"
/note="23778"
/complement(86..1204)
/function="orf; unknown function"
/note="Residues 1 to 372 of 372 are 100.00 pct identical to residues 1 to 372 of 372 from Escherichia coli K-12 strain MG1655: B2515"
/codon_start=1
/transl_table=1
/product="orf, hypothetical protein"
/protein_id="AAG57625.1"
/db_xref="GI:12516909"
/translation="MHNDAPIQRRKSTRIVGVNPIGDGAPIAVOSMTNPTTVEANVNQIKALERGADIVRVPTMDAEAFKIKQOVNPLVADIHFDRIALKAVEGYDCLRNPGNIGNERIRNVYDCARDKNIPRIQVNASLELDQEKYGEPPQALLESAMRHVDHLRLNFDQFKVSKASDVFLAVESYRLAKOIDPLHIGITEAGARGAVKSAIGLGLLEGGIDFLRVSLADPVEIKVGFILKSLRISRGINFACPSRCBEDVIGYNALGQLEDEITPMQVSIIGCVVNGEGALVSTLGVTGCKKSGLVEDGVRKDLNDNDMIDQLEARIKAKASQLEARRIDVOQVER"
/complement(1231..2244)
/gene="yfgA"
/note="23779"
/complement(1231..2244)
/gene="yfgA"
/note="23779"
/complement(1231..2244)
/function="putative membrane; Not classified"
/note="Residues 1 to 337 of 337 are 99.10 pct identical to residues 1 to 337 of 337 from Escherichia coli K-12 strain MG1655: B2516"
/codon_start=1
/transl_table=1
/product="putative membrane protein"
/protein_id="AAG57626.1"
/db_xref="GI:12516910"
/translation="MTEATHDQNEALTTGARLRNAREQLGSOQAVARLCLKVSTVRDIEDKADALASTFLRGYIRSVARLVHPEEKLPELEKQAPLRKAKVAPMOSFSLGRRKRRGDMFTMTLWLVFVIGLSGAMMDDHKAQOEEITTMADSSXELSSSEOGQSPYPLNTXTTDDRTSTTPASVDTTATNTQTPVTPAPAVDPOQNAVYSPSANYDPAATPATAATPDGAPLPDQAGVTTVPADPAALVMTADWELVTDATGKLLFGMQRKDNLM/LTGOAPYKLRIGAPAAVIOYQCKPDLRSRIRTNQVARRLLTNAEOSPAQ"
/complement(2529..3683)
/gene="yfgB"
/note="23780"
/complement(2529..3683)
/gene="yfgB"
/note="23780"
/function="orf; unknown function"
/note="Residues 1 to 384 of 384 are 99.47 pct identical to residues 1 to 384 of 384 from Escherichia coli K-12 strain MG1655: B2517"
/codon_start=1
/transl_table=1
/product="orf, hypothetical protein"
/protein_id="AAG57627.1"
/db_xref="GI:12516911"
/translation="MSBOLTPENTTKDKKINLIDLRQOMREFPKDGLGKPPRADOVAKMHTICDDEMDITNKVLKGLKEVAETRAPEVEERSSDGTIKMAIVAGDOVETVYIPEDDRATLCVSSQVGCALCECFSTAOQGFNRNLKRSSEIIGQVWAAATVGAATVGRPTINNVMMGEPLNLNNVPAAMEIMLDFGGLSKRRYTLSTSGVYPAADKQMDIDVALAISIHAPEIRDEIPIVINKXXETFLAVRYRLKRSNANOGRTVEYVWLHDVNGSTEHQOLAELLKDTPKKINIPNPAPAPYGRSSNRIDRFESKVLMSYGTITVIRKTRGDDIDACQQLAGVDYIDRKRLRRMOGEALIDIVAY"
/complement(3833..4264)
/gene="ndk"
/note="23781"
/complement(3833..4264)
/function="ndk"
/note="ndk"
/function="enzyme: Nucleotide biosynthesis: Purine ribonucleotide biosynthesis"
/note="Residues 1 to 143 of 143 are 100.00 pct identical

```

```

misc_feature
to residues 1 to 143 of 143 from Escherichia coli K-12 strain MG1655: B2518"
/codon_start=1
/transl_table=1
/product="nucleoside diphosphate kinase"
/protein_id="AAG57628.1"
/db_xref="GI:12516912"
/translation="MALERTFSIIRKPAVAKNVIQNTFARFEANGFKIVGKMLHLTVEOAGFYAEHDGKPEFDGLVEFMTSGPVIYSLGEGNAVORHDLGATNPANLACTLRADYADSLTENGHSDSVESAREIAIFYEGEGVCPRTT"
4397..9255
/note="O-island #106; Region of the EDL933 chromosome not homologous to E. coli K-12 MG1655"
/complement(4405..4983)
/gene="23782"
/complement(4405..4983)
/gene="23782"
/function="orf; unknown function"
/note="No significant matches"
/codon_start=1
/transl_table=1
/protein_id="AAG57629.1"
/db_xref="GI:12516913"
/translation="MLMHRREGITAPAIADLSQAKAMERALLINECLROYCEGMSVYSPHRTSEVAVSRRLHVQRADVQMTQIQPGLRLHTLMPQISDRPMPMDPOKCOLCGACWRACEQDOVESLNGHQLQINDALGCONCAVCFHOAMTVEITLPAKIVLNHANKVCKTCOKSFTEFQONAOQCLYCORHRRYGRMP"
/complement(5259..6080)
/gene="23783"
/complement(5259..6080)
/gene="23783"
/complement(5259..6080)
/function="putative enzyme; Not classified"
/note="Residues 2 to 272 of 273 are 25.26 pct identical to residues 6 to 276 of 286 from GenPept 118 : g1150021261gb1AADC7317.11AET35170.8 (AF135170) dimethyl sulfoxide reductase subunit C [Yersinia pestis]"
/codon_start=1
/transl_table=1
/product="putative dimethyl sulfoxide reductase subunit C"
/protein_id="AAG57630.1"
/db_xref="GI:12516914"
/translation="MHLEPLMAFTLLDQASVGMTLAFELRHILITSSAPPVPPPLLALAGAVGLLSTLHGLVPLNMAHLSHSSSLSEITFGALYTLALDLSTLLVALLRKGQMLMLAAVGVIVDVCMAQVYIHSITTWQYNTLIMFLGVGLGSLVAVFRISGILPQIDALRNGCVLIALVLRLLVOPLMIDLTANMAQATLDPHAPLMLGOLPILTLISGIVSGMFPFVAGCKKNIPALFGSVMLVSEVMLERFVFSIG"
/complement(6073..6510)
/gene="23784"
/complement(6073..6510)
/gene="23784"
/complement(6073..6510)
/function="putative enzyme; Not classified"
/note="Residues 4 to 134 of 145 are 61.83 pct identical to residues 57 to 186 of 205 from GenPept 118 : g1117871221gb1AAC73981.11(AE000191) anaerobic dimethyl sulfoxide reductase subunit B [Escherichia coli]"
/codon_start=1
/transl_table=1
/product="putative dimethyl sulfoxide reductase subunit B"
/protein_id="AAG57631.1"
/db_xref="GI:12516915"
/translation="MSHNVFATYTSVSCNHCADPCTKNCPTMMHMKRPGGIVAVNTDKVCGCGACSPGAPQOMEGQXMSKCFCDIDLQAKGQDPVAVCPGALIFGPIDLEKRYGSDVSKGLDPSITPNTLVIKPHOGAEKASHHA"
/complement(6699..9080)
/gene="23785"
/complement(6699..9080)
/gene="23785"
/complement(6699..9080)
/function="putative enzyme; Not classified"
/note="Residues 25 to 792 of 793 are 43.59 pct identical to residues 3 to 784 of 785 from GenPept 118 : g1117871221gb1AAC73980.11(AE000191) anaerobic dimethyl sulfoxide reductase subunit A [Escherichia coli]"
/codon_start=1

```

```

/transl_table=11
/product="putative dimethyl sulfoxide reductase subunit A"
/db_xref="GI:12516916"
/translation="MDPDEKERITPPQVRSRFLQATSLITLPISSTAKQSPDA
SPEVTAPADKVPVPTCSFDGSGCDIRAHMRGVVQIITLDPNEDLPMPIMRCV
RGKGYREYVHPRLKYPKMRVGRKREGKPERLSWEATRLIDNLKRIYQGPASR
YVHVGTVXGSGTSGDAMARRLNLGVEYHSLGATLAATPRTYQVAGSNM
DPLDRLVLLMGHNPTETFEHTNYFEFKMONGREFIYVDEPSDVSADONIP
LLEPTDMLADAMAVYISENLDKTEIDITYTGFPENSPRECVANESTVATLFAK
DSIHKTPEMAEKITHVPAOSIROLADYATTKRALIQSGPDRHICGERTAGSTLL
ASITGVNIGKGAAGTGGSSNRKFCVGPMPENPVQAKISIMNQADADASKVTPQ
DLKGVKDLSNIRLFLSLAGNYLANONPDVHQAALIDEDSEKIEFIVLSDLEMTSA
KYADVLIPETSGFMRNINIGETGCTASTYLISEKPIPEDFERFDYDMLRDVARKLVE
AEFSGRDEKQWIEHMEQTRILAMPDENLDPFATLQKTRHLKPSAPHIFEANIRDP
ONNPFTPSGKIEIFSKRLFEDMODPEPLASHVPAFEPEDGELTKKPYLOLITWTK
NRANSTOYAMPNLOEVOTOKLMLPNODAKORGISSEDSKVIYNDREVSTIPVITTRI
IPGVAMQAGAMQPDQGIIDRGCAVLSSTRITLAKONSHOTMIVEKERA"

```

Query Match 97.7%; Score 1093.4; DB 1; Length 11521;
 Best Local Similarity 98.6%; Pred. No. 3.1e-264;
 Matches 1103; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```

QY 1 ATGATAACCAAGGCTCCAAATTCAGTAGAAATCAACAGCATTTACGTTGGCAATGTG 60
D 1204 ATGATATACCAAGGCTCCAAATTCAGTAGAAATCAACAGCATTTACGTTGGCAATGTG 1145
QY 61 CCGATTGGCGATGGTGTCTCCATCGCCGTACAGTCATGACCAATACGCGTACGACAGAC 120
D 1144 CCGATTGGCGATGGTGTCTCCATCGCCGTACAGTCATGACCAATACGCGTACGACAGAT 1095
QY 121 GTCAACCAAGGTCATCAATCAACGAGCGGTGAGACGCGTGGCGTGAATATCGTCGT 180
D 1084 GTCAACCAAGGTCATCAATCAACGAGCGGTGAGACGCGTGGCGTGAATATCGTCGT 1025
QY 181 GTATCCGTACCGAGATGAGAGCGGCGAGAACGCTTCAACTCATCAACAGCAGCTTAAC 240
D 1024 GTATCCGTACCGAGATGAGAGCGGCGAGAACGCTTCAACTCATCAACAGCAGCTTAAC 965
QY 241 GTGCGCGTGTGGTGCATCCACTTCGACTATCGCATTCGCGTGAAGAGTACGGAATAC 300
D 964 GTGCGCGTGTGGTGCATCCACTTCGACTATCGCATTCGCGTGAAGAGTACGGAATAC 905
QY 301 GGGCGCATTTGTCTGCGTATTAACCCCTGGCAATATCGGTATGAAGAGCGTATTCGATG 360
D 904 GGGCGCATTTGTCTGCGTATTAACCCCTGGCAATATCGGTATGAAGAGCGTATTCGATG 845
QY 361 GTGGTTACTGTGGCGCGCATTAACATTCGATCCGTAFTGGCGCTTAACCGCGGATCG 420
D 844 GTGGTTACTGTGGCGCGCATTAACATTCGATCCGTAFTGGCGCTTAACCGCGGATCG 785
QY 421 CTGGAAGAAAGATCTGCAAGAAAGATATGCGAACCGACGCGGAGCGGTGCGAATCT 480
D 784 CTGGAAGAAAGATCTGCAAGAAAGATATGCGAACCGACGCGGAGCGGTGCGAATCT 725
QY 481 GCCATGGGTATGTATGATCATCTCGATCGCTGCAACTTCGATCAATCAAGTACGCGTG 540
D 724 GCCATGGGTATGTATGATCATCTCGATCGCTGCAACTTCGATCAATCAAGTACGCGTG 665
QY 541 AAAGCGCTGACGCTTCCGCGTGTGAGCTTATGCTTATGCGGCAAAACAGATCGAT 600
D 664 AAAGCGCTGACGCTTCCGCGTGTGAGCTTATGCTTATGCGGCAAAACAGATCGAT 605
QY 601 CACCGCTTGATCTGGGGATACCGAAGCGGTGTGTCGGCGACGCGGAGTAAATTC 660
D 604 CACCGCTTGATCTGGGGATACCGAAGCGGTGTGTCGGCGACGCGGAGTAAATTC 545
QY 661 GCCATTGGTTAGTGTCTGCTGTGAAGGATCGGGACACGCTGCGCGATACGCTG 720
D 544 GCCATTGGTTAGTGTCTGCTGTGAAGGATCGGGACACGCTGCGCGATACGCTG 485
QY 721 GCGGCGGATCCGCTGCAAGAGATCAAGTGCCTTGGATATTTTGAATCGCTGCGTATC 780
D GCGGCGGATCCGCTGCAAGAGATCAAGTGCCTTGGATATTTTGAATCGCTGCGTATC 1111

```

Db	484	GCGGCGGATCCGCTGCAAGAGATCAAGTGCCTTGGATATTTTGAATCGCTGATC	425
QY	781	CGTTCGGAGGATCAACTTCATCGCCCGCCAGCCTTTCGCTCAGGAATTTGATGT	840
Db	424	CGTTCGGAGGATCAACTTCATCGCCCGCCAGCCTTTCGCTCAGGAATTTGATGT	365
QY	841	ATCGTCAGGTTAAACGGCGTGGAGCAACGCGCTGGAATATCATCACTCCATGAGAGCTT	900
Db	364	ATCGTCAGGTTAAACGGCGTGGAGCAACGCGCTGGAATATCATCACTCCAGTACGCTT	305
QY	901	TGCATTTTCGCGTCCGTGCTGATGATGCGCCAGATGAGCGCTGTTTCTACACTCGCGCTC	960
Db	304	TGCATTTTCGCGTCCGTGCTGATGATGCGCCAGATGAGCGCTGTTTCTACACTCGCGCTC	245
QY	961	ACCGGCGGCAACGAAGAAAGCGGCTCTATGAAGATGGCGGCAAGAACCGCTGAGAC	1020
Db	244	ACCGGCGGCAACGAAGAAAGCGGCTCTATGAAGATGGCGGCAAGAACCGCTGAGAC	185
QY	1021	AACACGATATGATGACACGAGCTGGAAGACAGCATTCGTGCGAAGAGCAGCTGAGTGCAC	1080
Db	184	AACACGATATGATGACACGAGCTGGAAGACAGCATTCGTGCGAAGAGCAGCTGAGTGCAC	125
QY	1081	GAAGCGGCTGGAATTCAGCTTCACGAGCTTGAATAATTA 1119	
Db	124	GAAGCGGCTGGAATTCAGCTTCACGAGCTTGAATAATTA 86	
RESULT 9	AP002561/c	296827 bp	DNA linear BCT 07-MAR-2001
LOCUS	AP002561	296827 bp	DNA linear BCT 07-MAR-2001
DEFINITION	Escherichia coli O157:H7 DNA, complete genome, section 12/20.		
ACCESSION	AP002561	BA000007	
VERSION	AP002561.1	GI:13362592	
KEYWORDS	Escherichia coli O157:H7 (strain:O157:H7, sub_strain:RMD 0509952)		
SOURCE	DNA.		
ORGANISM	Escherichia coli O157:H7		
REFERENCE	Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.		
AUTHORS	1 (sites) Makino, K., Yokoyama, K., Kubota, Y., Yutsudo, C. H., Kimura, S., Kurokawa, K., Ishii, K., Hattori, M., Tatsuno, I., Abe, H., Iida, T., Yamamoto, K., Ohnishi, M., Hayashi, T., Yasunaga, T., Honda, T., Sasakawa, C., and Shingawa, H.		
TITLE	Complete nucleotide sequence of the prophage VP2-Sakai carrying the verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7 derived from the Sakai outbreak		
JOURNAL	Genes Genet. Syst. 74 (5), 227-239 (1999)		
MEDLINE	20196780		
REFERENCE	2 (sites) Ohnishi, M., Murata, T., Nakayama, K., Kuhara, S., Hattori, M., Kurokawa, K., Yasunaga, T., Yokoyama, K., Makino, K., Shingawa, H. and Hayashi, T.		
AUTHORS	Comparative analysis of the whole set of rRNA operons between an enterohemorrhagic Escherichia coli O157:H7 Sakai strain and an Escherichia coli K-12 strain M61655		
JOURNAL	Syst. Appl. Microbiol. 23 (3), 315-324 (2000)		
MEDLINE	20557356		
REFERENCE	3 (sites) Yokoyama, K., Makino, K., Kubota, Y., Watanabe, M., Kimura, S., Yutsudo, C. H., Kurokawa, K., Ishii, K., Hattori, M., Abe, H., Iida, T., Yamamoto, K., Hayashi, T., Yasunaga, T., Honda, T., Sasakawa, C. and Shingawa, H.		
AUTHORS	Complete nucleotide sequence of the prophage VP1-Sakai carrying the Shiga toxin 1 genes of the enterohemorrhagic Escherichia coli O157:H7 strain derived from the Sakai outbreak		
JOURNAL	Gene 258 (1-2), 127-139 (2000)		
MEDLINE	20564182		
REFERENCE	4 (sites) Hayashi, T., Makino, K., Ohnishi, M., Kurokawa, K., Ishii, K., Yokoyama, K., Han, C.-G., Ohtsubo, E., Nakayama, K., Murata, T., Tanaka, M., Tobe, T., Iida, T., Takami, H., Honda, T., Sasakawa, C., Ogasawara, N., Yasunaga, T., Kuhara, S., Shiba, T., Hattori, M. and		

TITLE
Shinagawa, H.
Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12

JOURNAL
DNA Res. 8 (1), 11-22 (2001)

REFERENCE
21156231

AUTHORS
5 (bases 1 to 296827)
Ohnishi, M., Kurokawa, K., Makino, K., Yasunaga, T., Shinagawa, H. and Hayashi, T.

TITLE
Direct Submission

JOURNAL
Submitted (26-JUN-2000) Ken Kurokawa, Osaka University, Genome Information Research Center, 3-1, Yamadaoka, Suita, Osaka 565-0871, Japan (E-mail: ken@gen-info.osaka-u.ac.jp),
URL: http://www.gen-info.osaka-u.ac.jp/, Tel: 81-6-6879-8365, Fax: 81-6-6879-2047)

COMMENT
genome project.

FEATURES
source
location/Qualifiers
1..296827
/organism="Escherichia coli O157:H7"
/strain="O157:H7"
/sub_strain="RIMD 0509952"
/db_xref="taxon:83334"
414..1484
/gene="ECS3123"
414..1484
/gene="ECS3123"
/note="unknown"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB36546.1"
/db_xref="GI:13362593"
/translation="MPNTHLSRCNLIQNNKLPTEIYKESQOAKLEPFCEDEFLKI
SOSRYGLSTDSIANLAFKASDAIDSIKTDIVHSGEYVIRHFEVHVIYKN
EPAAVEPYTLTFRGENEGILSVESKEGFSORTININEYTAIININEVIKENI
HNTVDSPTKISKINSSDKPKTKTSKESIKSGLSDRNIOCTYIITAKNTWYKKA
SSEERYDKFNKSKKEIKNTYKIQVOAKRLNOMKTYINKLSEGLDITDKKAPLINDS
TQEOYDKLIRYINIFPEPKLE"
complement(1708..2784)
/gene="ECS3124"
complement(1708..2784)
/gene="ECS3124"
/note="similar to GluP_ECOLI g11788572 percent identity
99 in 358 aa (Conserved in E.coli K-12)"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="glycerophosphodiester phosphodiesterase"
/protein_id="BAB36547.1"
/db_xref="GI:13362594"
/translation="MKLTNLNSMAIMNSTIVWGSSAMADSNKIKVIAHRGASGYLP
EHTLPKAMAVAGADYLEODLVMTKDDHLVYLHHYLDVATDVADRFEDRARKGRY
VAIDTLDLEIKLKTETEGFDIENGKKVQYPPRFPMGKSDFRVHFEETFEVQGLNH
STGKNGIYPEIKAPWFHQBGRDIAKTLVLRKTYTGKDKDYLYOCFTDDELKRI
KNELEPKMGMDLNLVQLIAYTDMNETQOKOPGSSVNNVYDMMKPGAKKQVAEYDAG
IGPDYHMLIEETSPGNIKLTGMVQDAQNKLVHPHYTVRSKLPETTTDQNLVAL
YKAGYNGLGFTDEPKAVFLNKE"
complement(2789..4147)
/gene="ECS3125"
complement(2789..4147)
/gene="ECS3125"
/note="similar to GluP_ECOLI g11788573 percent identity
99 in 452 aa (Conserved in E.coli K-12)"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="sn-glycerol-3-phosphate permease"
/protein_id="BAB36548.1"
/db_xref="GI:13362595"
/translation="MLSIKFPAPHKARLPAAEIDPYYRLRNOIFLGIFPGYAAYYLV
RKNEFLAMPYLVGSGRGLGALSGISIAFGFSKIVSGVSDNSNPFLPAGLIL
AAAVMLFPGFVPRATSSIAVMYVLLFLCGMFGMGMPGCGRTVMHWSKENGSIYSV

gene
CDS
4420..6048
/gene="ECS3126"
4420..6048
/gene="ECS3126"
/note="similar to GluP_ECOLI g11788574 percent identity
100 in 542 aa (Conserved in E.coli K-12)"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="anaerobic sn-glycerol-3-phosphate dehydrogenase
large subunit"
/protein_id="BAB36549.1"
/db_xref="GI:13362596"
/translation="MKTRDSQSDVYIIIGGATGAGIARDCAALRGVILVERDINT
GATGRNGLHSGARAVYTDASARECISENQILKRIARHCVEPTNGLFTLPEDDIS
FOATFIRACEEAGISABAIIDPOQARIIEPANNPALIGAVKPDGTVPFRLTANMD
AKEGAVILTAHEVYGLIREGATCGVARNHNLGEGOLALHAPVYVNAAGIMQHIIE
YADIRIRFPARKSLIIMDRINOHVINRCRPSDADIIVPGDTLSIGTSLRIDN
EIDNRVYAEVDILRLRGEKLAIVMAKTRILRASCYRPLVASDDPPSGRNVSGIV
LLDHAERGLDGFITITGKLMYRLMAEWYTDVACRLGVTCTTADLALPGSQEP
AEVTLRKVYISLPAPLRGSAVYRHGDRTPAMLSERLHSLVCEEAIVAGVQYAVEN
LNVSLDLRRRTYGMGTGCGELCAGCAALALDRFVNTTSAQISEQLSTPLNBRMG
VQPIAMGDALRESEFTRMVYGLGCGLEQKDAL"
6038..7297
/gene="ECS3127"
6038..7297
/gene="ECS3127"
/note="similar to GluP_ECOLI g11788575 percent identity
98 in 419 aa (Conserved in E.coli K-12)"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="anaerobic sn-glycerol-3-phosphate dehydrogenase
membrane anchor subunit"
/protein_id="BAB36550.1"
/db_xref="GI:13362597"
/translation="MRPDTIIMGGGLAGLLGLGLOKHLRCATVTRGOSALHPSGSS
IDLUSHPDQAVDIDHGLSRLRQAPAHRYSLGPRVVIDLACQALALAEAGAL
QGSVELPHQRTPLGTLSWLSSEVFWVWLPARKICVIGISIMDQAHAAASLR
BLDAVETAEIIELEPDLVRNNATEFRVNLARELDNEENPPLLDLALIPVANTCEMI
LMPACGLADDKLMRWLNELKPLCSLMLPTLPSPVYGLRLOQLORQVRRGGVWMPG
DEVKRVCKNGVYNEITRNHADIPLRPRFVLAGSFFSGIVLVERNGIREPILGLD
VLCATRGEMKKGDFEAPQPMQGEVTTDQTLRPSQAQGTIENLTIGSVLGGFPIA
QGGGCVAVSALHAQOIAQAGGQ"
7294..8484
/gene="ECS3128"
7294..8484
/gene="ECS3128"
/note="similar to GluP_ECOLI g11788576 percent identity
100 in 396 aa (Conserved in E.coli K-12)"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="anaerobic sn-glycerol-3-phosphate dehydrogenase
K small subunit"
/protein_id="BAB36551.1"
/db_xref="GI:13362598"
/translation="NNDVSFENCICVCTTACPYSRVNPGRPGKQKAPDGERLRK
DGALEADALKYCINCKRCEVAPDVKIGDIQARARAYDTTPSLRNEVLSHTLGM
SVSTFPAPIVNTATSLKPYRLDLAALIDHRRTLPKSCFGFRMYRSVAAQQAQYK
DOVAFPHGCFVYNNHOLGKDLIKVUNAMGTVOQLSKCCGVPILANGFIDKARKQ
ATNVESTREAVGKGPVIATRSCTFALRDEYEVNLVNDKGLRDIHELATRLMAR
KIDEGKTEPLKPLKAVYHTPCNHEKMGWLYTIELLRNIGLTLTVYDSCCCTIAG
TYGKREKNTPTSQAIGAPLFRQIESGADLVVTDCTCAQIEMSTSLRCEAPITLLA
QALA"
8484..9603
/gene="ECS3129"
8484..9603

CDS

8677..2603
/gene="Egs3129"
/note="Similar to YFAD_ECOLI g1178577 percent identity 94 in 308 aa but differs at C-ter (Conserved in E.coli K-12)"
/codon_start=1
/evidence="not_experimental
/translatable=1
/product="hypothetical protein"
/protein_id="BAB36552.1"

Query Match 97.7%; Score 1093.4; DB 1; Length 296827;
Best Local Similarity 96.0%; Pred. No. 3,7e-264;
Matches 1103; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGCATACCAAGGCTCCAAATTCACAGTAGAAATCAACACGATTTACTTGGGAATGTG 60
Db 280346 ATGCATACCAAGGCTCCAAATTCACAGTAGAAATCAACACGATTTACTTGGGAATGTG 280287

QY 61 CCGATGGCGATGGTGGTGGATGCGCGATGACATGATCCATATACCGGATGACAGAC 120
Db 280286 CCGATGGCGATGGTGGTGGATGCGCGATGACATGATCCATATACCGGATGACAGAT 280227

QY 121 GTCCAGCAAGGTCATGAATGAATCAAGGCGCTGAGACCGCTTGGCGCTGATATGCTCGT 180
Db 280226 GTCCAGCAAGGTCATGAATGAATCAAGGCGCTGAGACCGCTTGGCGCTGATATGCTCGT 280167

QY 181 GTATCCGTACCGAGATGAGCGCGAGAAAGCTTCAAACTCATCAACACAGAGTTTAC 240
Db 280166 GTATCCGTACCGAGATGAGCGCGAGAAAGCTTCAAACTCATCAACACAGAGTTTAC 280107

QY 241 GTGCGCGTGGTGGTGGATGACATGCTGACTATGCGATGGCTGTAAGAGTACGCGAATAC 300
Db 280106 GTGCGCGTGGTGGTGGATGACATGCTGACTATGCGATGGCTGTAAGAGTACGCGAATAC 280047

QY 301 GGCCTGCAATGCTGCGGATTAACCTGGCAATATCGGTATGAAGCGGTATCGCATG 360
Db 280046 GGCCTGCAATGCTGCGGATTAACCTGGCAATATCGGTATGAAGCGGTATCGCATG 279987

QY 361 GTGTTGACTGTGGCGCGATTAACATTCGATTCGATTTGGCTTAAACCGCGATGCG 420
Db 279986 GTGTTGACTGTGGCGCGATTAACATTCGATTCGATTTGGCTTAAACCGCGATGCG 279927

QY 421 CTGGAAGAAAGTCTGCAAGAAAGTATGCGGACCGGCGGCGGCGGAGTGAATCT 480
Db 279926 CTGGAAGAAAGTCTGCAAGAAAGTATGCGGACCGGCGGCGGCGGAGTGAATCT 279867

QY 481 GCCATGCGTATGTTGATCACTGATCGCTGAACTTGCATGATCAAGTCAAGCTG 540
Db 279866 GCCATGCGTATGTTGATCACTGATCGCTGAACTTGCATGATCAAGTCAAGCTG 279807

QY 541 AAAGCGTCAAGCTCTTCCGCTGTTGAGTCTTATGTTGCTGGCAAAACAGATCAT 600
Db 279806 AAAGCGTCAAGCTCTTCCGCTGTTGAGTCTTATGTTGCTGGCAAAACAGATCAT 279747

QY 601 CAGCGTGTGATCGGGGAGAGACGGAAGCGGCTGGCGGCGGCGGAGAGTAAATTC 660
Db 279746 CAGCGTGTGATCGGGGAGAGACGGAAGCGGCTGGCGGCGGCGGAGAGTAAATTC 279687

QY 661 GCCATTGTTAGTGTGCGCTGTCTGAAGGATCGGCGACACGCTGCCGATGCTG 720
Db 279686 GCCATTGTTAGTGTGCGCTGTCTGAAGGATCGGCGACACGCTGCCGATGCTG 279627

QY 721 GCGGCGATTCGCGTCCAGAGATCAAGTGGTGGATATTTGAAATGCTGCGATC 780
Db 279626 GCGGCGATTCGCGTCCAGAGATCAAGTGGTGGATATTTGAAATGCTGCGATC 279567

QY 781 CGTTCCGAGAGGATCAACTGATCGCTGCCAGACCGTGGCTGAGGAATTTGATGT 840
Db 279566 CGTTCCGAGAGGATCAACTGATCGCTGCCAGACCGTGGCTGAGGAATTTGATGT 279507

QY 841 ATCGGTACGGTATACGCGCTGAGCAAGCGCTGGAAGATATCATCTCGATGAGCAT 900
Db 279507 ATCGGTACGGTATACGCGCTGAGCAAGCGCTGGAAGATATCATCTCGATGAGCAT 279447

Db 279506 ATCGGTACGGTATACGCGCTGAGCAAGCGCTGGAAGATATCATCTCGATGAGCAT 279447

QY 901 TCGATTAATGCGCTCGCTGCTGTAATGCGCCAGCTGAGGCGCTGTTTCTACACTGCGGCTC 960
Db 279446 TCGATTAATGCGCTCGCTGCTGTAATGCGCCAGCTGAGGCGCTGTTTCTACACTGCGGCTC 279387

QY 961 ACCGCGGCAACAGAAAGAGCGGCTCTATGAAGATGGCGGCGCAAGACCGTCTGGAC 1020
Db 279386 ACCGCGGCAACAGAAAGAGCGGCTCTATGAAGATGGCGGCGCAAGACCGTCTGGAC 279327

QY 1021 AACACAGATATGATCGACACCTGGAAGCAGCATTCGTGCGAAAGCGCAGTCTGGAC 1080
Db 279326 AACACAGATATGATCGACACCTGGAAGCAGCATTCGTGCGAAAGCGCAGTCTGGAC 279267

QY 1081 GAAGCGGCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1119
Db 279266 GAAGCGGCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 279228

RESULT 10
AR027993 1010 bp DNA linear PAT 29-SEP-1999
LOCUS AR027993
DEFINITION Sequence 8 from patent US 5858367.
ACCESSION AR027993
VERSION AR027993.1 GI:5939966
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1010)
AUTHORS Rather,P.N.
TITLES Methods for screening for antimicrobials utilizing AarC and compositions thereof
JOURNAL Patent: US 5858367-A 8 12-JAN-1999;
FEATURES
source 1..1010
location/Qualifiers
BASE COUNT 238 a 253 c 287 g 232 t
ORIGIN

Query Match 90.3%; Score 1010; DB 6; Length 1010;
Best Local Similarity 100.0%; Pred. No. 2.6e-243;
Matches 1010; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCATACCAAGGCTCCAAATTCACAGTAGAAATCAACACGATTTACTTGGGAATGTG 60
Db 1 ATGCATACCAAGGCTCCAAATTCACAGTAGAAATCAACACGATTTACTTGGGAATGTG 60

QY 61 CCGATTGGCGATGGTGGTGGATGCGCGATGACATGATCCATATACCGGATGACAGAC 120
Db 61 CCGATTGGCGATGGTGGTGGATGCGCGATGACATGATCCATATACCGGATGACAGAC 120

QY 121 GTGGAAGCAAGGTCATCAATCAAGGCGCTGGAAGCGGCTGGCTGATATCGTCCGT 180
Db 121 GTGGAAGCAAGGTCATCAATCAAGGCGCTGGAAGCGGCTGGCTGATATCGTCCGT 180

QY 181 GTATCCGTACCGAGATGAGCGGCGAGAAAGCGTTCAAACTCATCAACAGCAGGTTTAA 240
Db 181 GTATCCGTACCGAGATGAGCGGCGAGAAAGCGTTCAAACTCATCAACAGCAGGTTTAA 240

QY 241 GTGCGCGTGGTGGCTGATCACTGACTGACTATGCGATGGCTGGAATGAGCGGATGAC 300
Db 241 GTGCGCGTGGTGGCTGATCACTGACTGACTATGCGATGGCTGGAATGAGCGGATGAC 300

QY 241 GTGCGCGTGGTGGCTGATCACTGACTGACTATGCGATGGCTGGAATGAGCGGATGAC 300
Db 241 GTGCGCGTGGTGGCTGATCACTGACTGACTATGCGATGGCTGGAATGAGCGGATGAC 300

QY 301 GCGGTGATTTGCTGCTGTTTAAACCTGGAATATGATGATGAAGAGCGTATGCGATG 360
Db 301 GCGGTGATTTGCTGCTGTTTAAACCTGGAATATGATGATGAAGAGCGTATGCGATG 360

QY 361 GTGCTTGAATGCTGCGCGCAATTAATTCAGTCCGATTCGATTTGGCTTAAACCGCGATG 420
Db 361 GTGCTTGAATGCTGCGCGCAATTAATTCAGTCCGATTCGATTTGGCTTAAACCGCGATG 420

QY 421 CTGGAAGAAAGTCTGCAAGAAAGTATGCGGACCGGCGGCGGCGGAGTGAATCT 480
Db 421 CTGGAAGAAAGTCTGCAAGAAAGTATGCGGACCGGCGGCGGCGGAGTGAATCT 480

Db	421	CTGGAAAAAGATCTCGAAGAAATATGGGAACCGACGGCCGAGGGTTGCTGGAAATCT	480
Qy	481	GCCATGCGTCATGTTGATCATCTCGATGCGCTGAACCTTGATCATGTTCAAAAGTCAGCGTG	540
Db	481	GCCATGCGTCATGTTGATCATCTCGATGCGCTGAACCTTGATCATGTTCAAAAGTCAGCGTG	540
Qy	541	AAACGCTGACGCTCCGCTCCGCGGTGAGCTTATGCTTATGCTTGGCGGAAAACGATGCGAT	600
Db	541	AAACGCTGACGCTCCGCTCCGCGGTGAGCTTATGCTTATGCTTGGCGGAAAACGATGCGAT	600
Qy	601	CAGCGCTTGATCTGAGGAGATCACCGAAGCGGTGTGCGCGACGCGGGCAGTAAATCC	660
Db	601	CAGCGCTTGATCTGAGGAGATCACCGAAGCGGTGTGCGCGACGCGGGCAGTAAATCC	660
Qy	661	GCCATTGGTTAGGTCGTGCTGCTCTCTGAAGGACATCGGCGACACGCTGCGCGTATGCGTG	720
Db	661	GCCATTGGTTAGGTCGTGCTGCTCTCTGAAGGACATCGGCGACACGCTGCGCGTATGCGTG	720
Qy	721	GGCGCGGATCCGGTCCGAAGATCAAAAGTGGTTTCGATATTTTGAATGCGTCGATC	780
Db	721	GGCGCGGATCCGGTCCGAAGATCAAAAGTGGTTTCGATATTTTGAATGCGTCGATC	780
Qy	781	CGTTTCGCGAGGAGATCACTTCATCGCTCCGCCGACCTGTTTCGCGTCAGAAATTTGATGT	840
Db	781	CGTTTCGCGAGGAGATCACTTCATCGCTCCGCCGACCTGTTTCGCGTCAGAAATTTGATGT	840
Qy	841	ATCGGTACGGTTAAACGGCGCTGAGCAACGCTGGAAATATCATCTCGATGAGAGCTT	900
Db	841	ATCGGTACGGTTAAACGGCGCTGAGCAACGCTGGAAATATCATCTCGATGAGAGCTT	900
Qy	901	TGCAATTATCGGCTCGTGTGTAAAGGCGCCAGGTGAGCGCGGTGTTCTACACTCGGCGTC	960
Db	901	TGCAATTATCGGCTCGTGTGTAAAGGCGCCAGGTGAGCGCGGTGTTCTACACTCGGCGTC	960
Qy	961	ACCGCGGCGCAACAAGAAAACGGCGCTCTATGTAAGATGCGTGGCGCAAGA 1010	
Db	961	ACCGCGGCGCAACAAGAAAACGGCGCTCTATGTAAGATGCGTGGCGCAAGA 1010	

RESULT	11								
AR208328									
LOCUS	AR208328		1010 bp	DNA					
DEFINITION	Sequence 8 from patent US 638745.								PAT 20-JUN-2002
ACCESSION	AR208328								
VERSION	AR208328.1		GI:21509454						
KEYWORDS:									
SOURCE	Unknown.								
ORGANISM	Unknown.								
REFERENCE	Unclassified.								
AUTHORS	1 (bases 1 to 1010)								
TITLE	Rather,P.N.								
JOURNAL	Method of screening for anti-microbial utilizing aarc and compositions thereof								
FEATURES	Patent: US 638745-A 8 07-MAY-2002;								
source	Location/Qualifiers								
	1..1010								
	/Organism="unknown"								
BASE COUNT	238 a	253 c	287 g	232 t					
ORIGIN									

[illegible]

OY	121	GTCCGAAGCAAGCGTAAATCAATCAAGCGCGTGAACCGCTGGGCGTAAATCGTCGGT	180
Db	121	GTCCGAAGCAAGCGTAAATCAATCAAGCGCGTGAACCGCGTGGGCGTAAATCGTCGGT	180
OY	181	GTATCCGTACCGAGATGGAGCGGGCAGAGCGCTTCAAACTCATCAACAGCAGGTTAAC	240
Db	181	GTATCCGTACCGAGATGGAGCGGGCAGAGCGCTTCAAACTCATCAACAGCAGGTTAAC	240
OY	241	GTGCGCGTGGTGGCTGCATCCACTTTCGACATATGCGCTGAAGTACCGGAATAC	300
Db	241	GTGCGCGTGGTGGCGACATCCACTTTCGACATATGCGCTGAAGTAAAGTAAAGTAC	300
OY	301	GCGCGTGAATGCTGCGATTAAACCTGGCAATATCGGTTAATGAAGACGTAATTTGCATG	360
Db	301	GCGCGTGAATGCTGCGATTAAACCTGGCAATATCGGTTAATGAAGACGTAATTTGCATG	360
OY	361	GTGTTGACTGTGGCGCGATTAATAAACAATTCGATCCGTAATTTGGCCGTTAACGCCGATC	420
Db	361	GTGTTGACTGTGGCGCGATTAATAAACAATTCGATCCGTAATTTGGCCGTTAACGCCGATC	420
OY	421	CTGGAAAAAGATCTGCAAGAAAAGTATGGCGAACCGCAGCGCGTTCGTGGAATCT	480
Db	421	CTGGAAAAAGATCTGCAAGAAAAGTATGGCGAACCGCAGCGCGTTCGTGGAATCT	480
OY	481	GCCATGCGTCAATGTTGATCATCTCGATCGCCGTAACCTTGATCAATCAAGTCAGCGTG	540
Db	481	GCCATGCGTCAATGTTGATCATCTCGATCGCCGTAACCTTGATCAATCAAGTCAGCGTG	540
OY	541	AAAGCGTCTGACGCTTCTCTCGCTGTGAGTCTTATCGTTTCTCTGGCAAAACAGTTCAT	600
Db	541	AAAGCGTCTGACGCTTCTCTCGCTGTGAGTCTTATCGTTTCTCTGGCAAAACAGTTCAT	600
OY	601	CAGCGGTCATCTGGGCGATCACCGAAGCGCGGTGGTGGCGCGCAGCGGGCAGTAATAATCC	660
Db	601	CAGCGGTCATCTGGGCGATCACCGAAGCGCGGTGGTGGCGCGCAGCGGGCAGTAATAATCC	660
OY	661	GCCATTGGTTTAGTCTGCTGCTGTCTGAAGGCATCGGCGACACGCTGGCGCTATCCGCTG	720
Db	661	GCCATTGGTTTAGTCTGCTGCTGTCTGAAGGCATCGGCGACACGCTGGCGCTATCCGCTG	720
OY	721	GCGGCGGATCCGAGTGAAGAGATCAAAAGCGGTTTCGATATTTTGAATAAATCGCTGCATAC	780
Db	721	GCGGCGGATCCGAGTGAAGAGATCAAAAGCGGTTTCGATATTTTGAATAAATCGCTGCATAC	780
OY	781	CGTTGCGAGGAGATCACTTCATCGCGCTGGCCGACCTGTTGCGGTGAGGAATTTGATGTT	840
Db	781	CGTTGCGAGGAGATCACTTCATCGCGCTGGCCGACCTGTTGCGGTGAGGAATTTGATGTT	840
OY	841	ATCGGTACGCTTAACGCGTGGAGCAACGCCCTGGAAAGATATCATCTCCATGACGCTT	900
Db	841	ATCGGTACGCTTAACGCGTGGAGCAACGCCCTGGAAAGATATCATCTCCATGACGCTT	900
OY	901	TGCAATTATGCGCTGGTGTGAATGGCCAGGTGAGGCGCTGTTTCTACACTCGGCGCTC	960
Db	901	TGCAATTATGCGCTGGTGTGAATGGCCAGGTGAGGCGCTGTTTCTACACTCGGCGCTC	960
OY	961	ACCGCGGCAACAGAAAAGCGGCTCTATGAAGATGGCTGCGCCACAAGA 1010	
Db	961	ACCGCGGCAACAGAAAAGCGGCTCTATGAAGATGGCTGCGCCACAAGA 1010	

RESULT 12	AE008814/c	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
	AE008814	23647 bp	DNA, linear	Salmonella typhimurium LT2, section 118 of 220 of the complete genome.	AE008814.1	GI:16421058	Salmonella typhimurium LT2.	Salmonella typhimurium LT2.
							Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Salmonella.	

REFERENCE 1 (bases 1 to 23647)
 AUTHORS McCelland, M., Sanderson, K.E., Spieth, J., Clifton, S.W.,
 Latreille, P., Courtney, L., Portolick, S., Ali, J., Dante, S.W.,
 Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A.,
 Grewal, N., Mulvaney, E., Ryan, E., Sun, H., Flores, L., Miller, W.,
 Steneking, T., Nham, M., Waterston, R., and Wilson, R.K.
 TITLE Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 LT2
 JOURNAL Nature 413 (6858) 852-856 (2001)
 MEDLINE 21534948
 PUBMED 11677609
 REFERENCE 2 (bases 1 to 23647)
 AUTHORS The Salmonella typhimurium Genome Sequencing Project.
 TITLE Direct Submission
 JOURNAL Submitted (29-MAR-2001) Genome Sequencing Center, Department of
 Genetics, Washington University School of Medicine, 4444 Forest
 Park Boulevard, St. Louis, MO 63108, USA
 COMMENT Supported by NIH grant 5U 01 AI43283

COMMENT Coding sequences below are predicted from manually evaluated
 computer analysis using similarity information and the programs:
 GLIMMER: <http://www.tigr.org/softlab/glimmer/glimmer.html> and
 GeneMark: <http://opal.biology.gatech.edu/GeneMark/>

EC numbers were kindly provided by Junko Yabuzaki and the Kyoto
 Encyclopedia of Genes and Genomes: <http://www.genome.ad.jp/kegg/>,
 and Pedro Romero and Peter Karp at EcoCyc:
<http://ecocyc.pangeasystems.com/ecocyc/>

The analyses of ribosome binding sites and promoter binding sites
 were kindly provided by Heladia Salgado, Julio Collado-Vides and
 Reguondb:
http://kinch.cifn.unam.mx:8850/db/reguondb_intro.frameset

This sequence was finished as follows unless otherwise noted: all
 regions were double stranded, sequenced with an alternate
 chemistries or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one m3 subclone.

FEATURES
 source
 location/qualifiers
 1..23647
 /organism="Salmonella typhimurium LT2"
 /strain="9812; SGSC 1412; ATCC 700720"
 /db_xref="ATCC:700720"
 /db_xref="taxon:99287"
 /note="LT2"
 complement(110..5711)
 /gene="rfaA"
 /note="rfaA (g14583529)"
 /codon_start=1
 /transl_table=11
 /product="putative outer membrane protein"
 /protein_id="AAL21409.1"
 /db_xref="GI:16421059"
 /translation="MDNRNKKVKKVHFLTLIMFVSTAPPAALLKGTGNOELNSVT
 GAVNGAPRLDGAIIPLRYOGSTLIDSKTHDIEFSSAPROFSDATSTSMRANSTDT
 EGDLSDFZLIAMENKOPPAKGLVWADATPTPLSPQVNLITFCQWLAKAROLAW
 AOEDETNYEALMLFTRIGVPYNAIPLSPYVALNIKPATDTPVYSDHDAFEA
 SKVGESELTISTKACDGTPEGNAPFVIRKDAENRGVNNAPVAGDTELT
 OMEYRGVANGNATVYVTKEGPGKPLVASSVNEPALTAETAVIEFTTISPSDCK
 ASMGHMTSATATNGITVYFTRPLIAEASGADKSVDTMETALFTWGSADNHC
 ILPDAEOLYMHHESTLETTGWGWPAGAEYWSSTKNOISGHAHVHNSASVRAA
 DSDTLVSGYKAQPAHPQITLSDEGPKAAGVSTIDLTMTVDRDQKPLPYRME
 LEIDPAKNGEHODAMNLRVTVDSEGSASSPEHYTGTVDVNGAHLTLKHSGLG
 VETPIRIYEDDEGNVELSFVETVYSPVDGAMNMGHGVVADAGNLKRLPLA
 VASHKDGSEENEMENATENSASATACGAGVGDOSLSAHLVSEHAGGOMESPHG
 WPTEDYFIADSDASGTIVNLENGDSGFTPTPTLTSANEMAVLVDYVINDPVI
 KNAADKVGSEGVKANHSHRNALMGALIGNTDPTITNANGRRDGLTTGFTDVSXKMF
 DDVGYVAGGVHGTIDANGDATIILTKKGVLPLPLNIAPVDSLITPVSXKFTV

gene
 complement(110..5707)
 /gene="rfaA"
 /note="rfaA (g14583529)"
 /codon_start=1
 /transl_table=11
 /product="putative outer membrane protein"
 /protein_id="AAL21409.1"
 /db_xref="GI:16421059"
 /translation="MDNRNKKVKKVHFLTLIMFVSTAPPAALLKGTGNOELNSVT
 GAVNGAPRLDGAIIPLRYOGSTLIDSKTHDIEFSSAPROFSDATSTSMRANSTDT
 EGDLSDFZLIAMENKOPPAKGLVWADATPTPLSPQVNLITFCQWLAKAROLAW
 AOEDETNYEALMLFTRIGVPYNAIPLSPYVALNIKPATDTPVYSDHDAFEA
 SKVGESELTISTKACDGTPEGNAPFVIRKDAENRGVNNAPVAGDTELT
 OMEYRGVANGNATVYVTKEGPGKPLVASSVNEPALTAETAVIEFTTISPSDCK
 ASMGHMTSATATNGITVYFTRPLIAEASGADKSVDTMETALFTWGSADNHC
 ILPDAEOLYMHHESTLETTGWGWPAGAEYWSSTKNOISGHAHVHNSASVRAA
 DSDTLVSGYKAQPAHPQITLSDEGPKAAGVSTIDLTMTVDRDQKPLPYRME
 LEIDPAKNGEHODAMNLRVTVDSEGSASSPEHYTGTVDVNGAHLTLKHSGLG
 VETPIRIYEDDEGNVELSFVETVYSPVDGAMNMGHGVVADAGNLKRLPLA
 VASHKDGSEENEMENATENSASATACGAGVGDOSLSAHLVSEHAGGOMESPHG
 WPTEDYFIADSDASGTIVNLENGDSGFTPTPTLTSANEMAVLVDYVINDPVI
 KNAADKVGSEGVKANHSHRNALMGALIGNTDPTITNANGRRDGLTTGFTDVSXKMF
 DDVGYVAGGVHGTIDANGDATIILTKKGVLPLPLNIAPVDSLITPVSXKFTV

CDS
 complement(110..5707)
 /gene="rfaA"
 /note="rfaA (g14583529)"
 /codon_start=1
 /transl_table=11
 /product="putative outer membrane protein"
 /protein_id="AAL21409.1"
 /db_xref="GI:16421059"
 /translation="MDNRNKKVKKVHFLTLIMFVSTAPPAALLKGTGNOELNSVT
 GAVNGAPRLDGAIIPLRYOGSTLIDSKTHDIEFSSAPROFSDATSTSMRANSTDT
 EGDLSDFZLIAMENKOPPAKGLVWADATPTPLSPQVNLITFCQWLAKAROLAW
 AOEDETNYEALMLFTRIGVPYNAIPLSPYVALNIKPATDTPVYSDHDAFEA
 SKVGESELTISTKACDGTPEGNAPFVIRKDAENRGVNNAPVAGDTELT
 OMEYRGVANGNATVYVTKEGPGKPLVASSVNEPALTAETAVIEFTTISPSDCK
 ASMGHMTSATATNGITVYFTRPLIAEASGADKSVDTMETALFTWGSADNHC
 ILPDAEOLYMHHESTLETTGWGWPAGAEYWSSTKNOISGHAHVHNSASVRAA
 DSDTLVSGYKAQPAHPQITLSDEGPKAAGVSTIDLTMTVDRDQKPLPYRME
 LEIDPAKNGEHODAMNLRVTVDSEGSASSPEHYTGTVDVNGAHLTLKHSGLG
 VETPIRIYEDDEGNVELSFVETVYSPVDGAMNMGHGVVADAGNLKRLPLA
 VASHKDGSEENEMENATENSASATACGAGVGDOSLSAHLVSEHAGGOMESPHG
 WPTEDYFIADSDASGTIVNLENGDSGFTPTPTLTSANEMAVLVDYVINDPVI
 KNAADKVGSEGVKANHSHRNALMGALIGNTDPTITNANGRRDGLTTGFTDVSXKMF
 DDVGYVAGGVHGTIDANGDATIILTKKGVLPLPLNIAPVDSLITPVSXKFTV

ATSPDPAKMGHMDITVGDWTFERPLAGEVNSPLRTODESNETVARVHSDAV
 GNDGAGCANRLPRIIDLEALYANNSGKIHSGOMPTDYNWSTYOSATWKLIA
 LTNSEFANSNPIYDASCLASDNPVAVASITIEVNSQWTDGSDVAVVKGKGTQOL
 KTVKQVSGPIPEAPFVLTRGDGIDRGRKRTAOUQDDGDLVTVVLDGESLATT
 RKMGGDTGIDGTRILISVTRPDHGTFRALNATLYENAAVASIDTLETVYSDSVAR
 RMGNHAPSLTLADGAVOYRPLLYAEIASSDNDKMAQDEISLKGSTLYMOYKDKTKGN
 YMSENDNGYVLDCLTTSRSGNLIALSDANADKSAVAKGETIIPMTVYTNAG
 OPQAGVAVLTRDYAYSRGAVDKOYIEPGVIGEPVPTSPANMMLAPVAPGTAFAF
 MNONGSTKWSGCTGDDKRLPTLTDKSLGKTSTALANOPEAASVDAFTYOT
 SPDTPIASTWGHMPTVOYNGVTLRPTIKAEISAPRDTWPNNEFWGTNTYOSSEH
 VETSLHLCGSDENIASLIDLKALQSVIGTLQWPTTSDWVYSDGSKNKKYCSFNE
 TTGQTTCTREKATTSGLGCRVP"
 complement(5706..5711)
 /gene="rfaA"
 /note="putative RBS for rfaA; Reguondb: STMSIH002622"
 complement(5827..6793)
 /gene="sinI"
 /note="STM2516"
 complement(5827..6786)
 /gene="sinI"
 /note="sinI (g14583530)"
 /codon_start=1
 /transl_table=11
 /product="putative outer membrane protein"
 /protein_id="AAL21410.1"
 /db_xref="GI:16421060"
 /translation="MQATVKKRLTFLKALAVVAGYCAAPAVANGKNGKSGMOIVSE
 TGTIOCTVWTRAAKRTADTDKDHVTYIDGDKRTYEGSKQFHVGRKTYVNAIG
 DTBGDLDTNATKATKAVNSFSQDNGSPKIDGSDSTEIOAADDRITGKLTPTT
 TTGDPVAVATELLKIDSLDAGGSDDEIPEBYPDENVHVYIHKDSNTNLKNSGT
 TUKNTYOVLLMSDKNGSTGADAGENVDQDYRKRFVGTGKINGTGTGIVNESWN
 DKDILVPTVNAAKRTAFEGADGVGVGSDGVGFGISIDYKRR"
 complement(6786..6793)
 /gene="sinI"
 /note="putative RBS for sinI; Reguondb: STMSIH002623"
 complement(6844..9042)
 /gene="sinH"
 /note="STM2517"
 complement(6844..9036)
 /gene="sinH"
 /note="similar to Escherichia coli intimin and Yersinia
 pestis invasins proteins; (g14583531)"
 /codon_start=1
 /transl_table=11
 /product="sinH"
 /protein_id="AAL21411.1"
 /db_xref="GI:16421061"
 /translation="WVRIRILVLLVSGVYVAAAGDNGSASANDSHSSLPDLASES
 ARKEDEKNGKSEKGEADYFINSATQGDNLTPALBSQARSYLQDITSSQSYLE
 GVMSPYKIRTSLSVGEGLDSSLDYIIPYDQSGITFSQISAKREKEDITNFG
 LGVONVGMWLLGNAFYDYDTPRGHRLGITEAMTVDLKSQVNYHPLSPKMSKD
 PDYEFERPARMGDIMESEWLPRYOLGKIVYEGVEVALFGTDNLQKDHAWTIG
 LEYTVPLVTVTDKAKGSDNSVYATVYQICTPLAADIDPENKILQHSLSGR
 TDFVDNNFTILEYREKDPDVTLLWUKADATNEHPECVIEDPEAAGLEKRWYNA
 LTNHKKIISASWQAKNNAARTLVAVYKANAULTGNNNSMWLVPAWNAUTEBQRT
 ALNTKRVKTLDEDEKGNKONGSVVEITVOODKRLIYDNIADTDRSDSHSASALAD
 GEDGVVMDLLTSDGSDSDRNGENLVDAMPVLYSDSKVTLTAQDPTCTETPSCV
 IASRKEAGTVVLTSLTSGTFEWRKAKADAGVQVVDVFTGDNLSALNAVYQVAA
 NQVNLIGKDKRPYNNAYRILMDKDKKQGVFGQSGDTEBEALYLYQMEFTQST
 NGHTGALANTNMEDLVLPVTNKAQKFAANVEDSGVGIRVITYSOK"
 complement(9037..9042)
 /gene="sinH"
 /note="putative RBS for sinH; Reguondb: STMSIH002624"
 complement(9382..9615)
 /gene="yfgJ"
 /note="STM2518"
 complement(9382..9603)

RBS
 complement(5706..5711)
 /gene="rfaA"
 /note="putative RBS for rfaA; Reguondb: STMSIH002622"
 complement(5827..6793)
 /gene="sinI"
 /note="STM2516"
 complement(5827..6786)
 /gene="sinI"
 /note="sinI (g14583530)"
 /codon_start=1
 /transl_table=11
 /product="putative outer membrane protein"
 /protein_id="AAL21410.1"
 /db_xref="GI:16421060"
 /translation="MQATVKKRLTFLKALAVVAGYCAAPAVANGKNGKSGMOIVSE
 TGTIOCTVWTRAAKRTADTDKDHVTYIDGDKRTYEGSKQFHVGRKTYVNAIG
 DTBGDLDTNATKATKAVNSFSQDNGSPKIDGSDSTEIOAADDRITGKLTPTT
 TTGDPVAVATELLKIDSLDAGGSDDEIPEBYPDENVHVYIHKDSNTNLKNSGT
 TUKNTYOVLLMSDKNGSTGADAGENVDQDYRKRFVGTGKINGTGTGIVNESWN
 DKDILVPTVNAAKRTAFEGADGVGVGSDGVGFGISIDYKRR"
 complement(6786..6793)
 /gene="sinI"
 /note="putative RBS for sinI; Reguondb: STMSIH002623"
 complement(6844..9042)
 /gene="sinH"
 /note="STM2517"
 complement(6844..9036)
 /gene="sinH"
 /note="similar to Escherichia coli intimin and Yersinia
 pestis invasins proteins; (g14583531)"
 /codon_start=1
 /transl_table=11
 /product="sinH"
 /protein_id="AAL21411.1"
 /db_xref="GI:16421061"
 /translation="WVRIRILVLLVSGVYVAAAGDNGSASANDSHSSLPDLASES
 ARKEDEKNGKSEKGEADYFINSATQGDNLTPALBSQARSYLQDITSSQSYLE
 GVMSPYKIRTSLSVGEGLDSSLDYIIPYDQSGITFSQISAKREKEDITNFG
 LGVONVGMWLLGNAFYDYDTPRGHRLGITEAMTVDLKSQVNYHPLSPKMSKD
 PDYEFERPARMGDIMESEWLPRYOLGKIVYEGVEVALFGTDNLQKDHAWTIG
 LEYTVPLVTVTDKAKGSDNSVYATVYQICTPLAADIDPENKILQHSLSGR
 TDFVDNNFTILEYREKDPDVTLLWUKADATNEHPECVIEDPEAAGLEKRWYNA
 LTNHKKIISASWQAKNNAARTLVAVYKANAULTGNNNSMWLVPAWNAUTEBQRT
 ALNTKRVKTLDEDEKGNKONGSVVEITVOODKRLIYDNIADTDRSDSHSASALAD
 GEDGVVMDLLTSDGSDSDRNGENLVDAMPVLYSDSKVTLTAQDPTCTETPSCV
 IASRKEAGTVVLTSLTSGTFEWRKAKADAGVQVVDVFTGDNLSALNAVYQVAA
 NQVNLIGKDKRPYNNAYRILMDKDKKQGVFGQSGDTEBEALYLYQMEFTQST
 NGHTGALANTNMEDLVLPVTNKAQKFAANVEDSGVGIRVITYSOK"
 complement(9037..9042)
 /gene="sinH"
 /note="putative RBS for sinH; Reguondb: STMSIH002624"
 complement(9382..9615)
 /gene="yfgJ"
 /note="STM2518"
 complement(9382..9603)

gene
 complement(110..5707)
 /gene="rfaA"
 /note="rfaA (g14583529)"
 /codon_start=1
 /transl_table=11
 /product="putative outer membrane protein"
 /protein_id="AAL21409.1"
 /db_xref="GI:16421059"
 /translation="MDNRNKKVKKVHFLTLIMFVSTAPPAALLKGTGNOELNSVT
 GAVNGAPRLDGAIIPLRYOGSTLIDSKTHDIEFSSAPROFSDATSTSMRANSTDT
 EGDLSDFZLIAMENKOPPAKGLVWADATPTPLSPQVNLITFCQWLAKAROLAW
 AOEDETNYEALMLFTRIGVPYNAIPLSPYVALNIKPATDTPVYSDHDAFEA
 SKVGESELTISTKACDGTPEGNAPFVIRKDAENRGVNNAPVAGDTELT
 OMEYRGVANGNATVYVTKEGPGKPLVASSVNEPALTAETAVIEFTTISPSDCK
 ASMGHMTSATATNGITVYFTRPLIAEASGADKSVDTMETALFTWGSADNHC
 ILPDAEOLYMHHESTLETTGWGWPAGAEYWSSTKNOISGHAHVHNSASVRAA
 DSDTLVSGYKAQPAHPQITLSDEGPKAAGVSTIDLTMTVDRDQKPLPYRME
 LEIDPAKNGEHODAMNLRVTVDSEGSASSPEHYTGTVDVNGAHLTLKHSGLG
 VETPIRIYEDDEGNVELSFVETVYSPVDGAMNMGHGVVADAGNLKRLPLA
 VASHKDGSEENEMENATENSASATACGAGVGDOSLSAHLVSEHAGGOMESPHG
 WPTEDYFIADSDASGTIVNLENGDSGFTPTPTLTSANEMAVLVDYVINDPVI
 KNAADKVGSEGVKANHSHRNALMGALIGNTDPTITNANGRRDGLTTGFTDVSXKMF
 DDVGYVAGGVHGTIDANGDATIILTKKGVLPLPLNIAPVDSLITPVSXKFTV

CDS
 complement(110..5707)
 /gene="rfaA"
 /note="rfaA (g14583529)"
 /codon_start=1
 /transl_table=11
 /product="putative outer membrane protein"
 /protein_id="AAL21409.1"
 /db_xref="GI:16421059"
 /translation="MDNRNKKVKKVHFLTLIMFVSTAPPAALLKGTGNOELNSVT
 GAVNGAPRLDGAIIPLRYOGSTLIDSKTHDIEFSSAPROFSDATSTSMRANSTDT
 EGDLSDFZLIAMENKOPPAKGLVWADATPTPLSPQVNLITFCQWLAKAROLAW
 AOEDETNYEALMLFTRIGVPYNAIPLSPYVALNIKPATDTPVYSDHDAFEA
 SKVGESELTISTKACDGTPEGNAPFVIRKDAENRGVNNAPVAGDTELT
 OMEYRGVANGNATVYVTKEGPGKPLVASSVNEPALTAETAVIEFTTISPSDCK
 ASMGHMTSATATNGITVYFTRPLIAEASGADKSVDTMETALFTWGSADNHC
 ILPDAEOLYMHHESTLETTGWGWPAGAEYWSSTKNOISGHAHVHNSASVRAA
 DSDTLVSGYKAQPAHPQITLSDEGPKAAGVSTIDLTMTVDRDQKPLPYRME
 LEIDPAKNGEHODAMNLRVTVDSEGSASSPEHYTGTVDVNGAHLTLKHSGLG
 VETPIRIYEDDEGNVELSFVETVYSPVDGAMNMGHGVVADAGNLKRLPLA
 VASHKDGSEENEMENATENSASATACGAGVGDOSLSAHLVSEHAGGOMESPHG
 WPTEDYFIADSDASGTIVNLENGDSGFTPTPTLTSANEMAVLVDYVINDPVI
 KNAADKVGSEGVKANHSHRNALMGALIGNTDPTITNANGRRDGLTTGFTDVSXKMF
 DDVGYVAGGVHGTIDANGDATIILTKKGVLPLPLNIAPVDSLITPVSXKFTV

```

/gene="yfgu"
/note="similar to E. coli orf, hypothetical protein
(AAC75563.1); Blastp hit to AAC75563.1 (83 aa), 63%
identity in aa 13 - 83"
/codon_start=1
/transl_table=11
/product="putative cytoplasmic protein"
/protein_id="AAL21412.1"
/db_xref="GI:16421062"
/translation="MEINCPVCHALEHRENSPTAHCETCAKDFSLQALCPDORPQVLY
KACGAVDFECNGHGLSKRKNVVISQ"
RBS
/gene="yfgu"
/note="putative RBS for yfgu; RegulonDB:STMSI002625"
/complement(9692..11176)
/gene="engA"
/note="STM519"
/complement(9692..11164)
/gene="engA"
/note="similar to E. coli putative GTP-binding factor
(AAC75564.1); Blastp hit to AAC75564.1 (503 aa), 97%
identity in aa 14 - 503"
/codon_start=1
/transl_table=11
/product="putative GTP-binding protein"
/protein_id="AAL21413.1"
/db_xref="GI:16421063"
/transl_table=11
/note="putative RBS for engA; RegulonDB:STMSI002626"
/complement(11283..12471)
/gene="yfgL"
/note="STM520"
/complement(11283..12461)
/gene="yfgL"
CDS
Query Match 79.3% Score 887; DB 1; Length 23647;
Best Local Similarity 87.0%; Pred. No. 2.8e-212;
Matches 974; Conservative 0; Mismatches 145; Indels 0; Gaps 0;
QY 1 ATGCATTAACGAGCTCCATTCAAGTAGAATAATCAACACGATTTACGTTGGGAATGTG 60
15609 ATGCATTAACGAGCTCCATTCAAGTAGAATAATCAACACGATTTACGTTGGGAATGTG 15550
QY 61 CCATTGGCGATGTGCTCCATCGCCGCTACATGACATGACCAATACGCTAGACAGAC 120
15549 CCATTGGCGATGTGCTCCATCGCCGCTACATGACATGACCAATACGCTAGACAGAC 15490
QY 121 GTGGAAGCAACGGTCATCAATCAAGGCGCTGGAACGCGTTGGCGCTGATCGTCGCT 180
15489 GTGGAAGCAACGGTCATCAATCAAGGCGCTGGAACGCGTTGGCGCTGATCGTCGCT 15430
QY 181 GTATCCGTACCGAGATGAGCGGCGGAGAGCGTTCAAACTCAAAACAGAGTTTAC 240
15429 GTTCTGTACCGAANTGAGCGGCGGAGAGCGTTCAAACTCAAAACAGAGTTTAC 15370
QY 241 GTGCGCGTGTGCTGATCAATCCATTGATGATGCGCTTGGCGTGAAGAGGGAATAC 300
15369 GTGCGCGTGTGCTGATCAATCCATTGATGATGCGCTTGGCGTGAAGAGGGAATAC 15310
QY 301 GGGCGATTTGCTGGCTATTAAACCCGTGGCAATATCGGTATGAGAGCGGATTCGCGATG 360
15309 GGGCGATTTGCTGGCTATTAAACCCGTGGCAATATCGGTATGAGAGCGGATTCGCGATG 15250

```

```

QY 361 GTGTTGACTGTGCGCGGAGTAATAACATTCGATCCGATTTGGCGTTAACCCGAGTC 420
15249 GTGTTGACTGTGCGCGGAGTAATAACATTCGATCCGATTTGGCGTTAACCCGAGTC 15190
QY 421 CTGGAATAAAGATCTGCAAGAAAGATGAGCGGAGAGCGCGGAGCGGCTGCTGGAATCT 480
15189 CTGGAATAAAGATCTGCAAGAAAGATGAGCGGAGAGCGCGGAGCGGCTGCTGGAATCT 15130
QY 481 GCCATGCGTCATCTGTGATCATCTCGATGCGCCGTAACCTTGATCATCAAGTACGCGTG 540
15129 GCCATGCGTCATCTGTGATCATCTCGATGCGCCGTAACCTTGATCATCAAGTACGCGTG 15070
QY 541 AAAGCTGTGACGCTTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
15069 AAAGCTGTGACGCTTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 15010
QY 601 CAGCGGTTCATCTGCGGAGTACCGAGAGCGGTGCGCGGAGCGGCGGAGCAATAATCC 660
15009 CAGCGGTTCATCTGCGGAGTACCGAGAGCGGTGCGCGGAGCGGCGGAGCAATAATCC 14950
QY 661 GCCATTTGTTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
14949 GCCATTTGTTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 14890
QY 721 GCGGCGGATCCGCTGGAAGATCAACAGTGGTTTCGATTTTGAATCCGTCGCTATC 780
14889 GCGGCGGATCCGCTGGAAGATCAACAGTGGTTTCGATTTTGAATCCGTCGCTATC 14830
QY 781 CGTTCGCGAGGATCACTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
14829 CGTTCGCGAGGATCACTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14770
QY 841 ATCGGATCGGTTAAGCGCTGAGAGCAACGCTGGAAGATATCATCTCCATGACGTT 900
14769 ATCGGATCGGTTAAGCGCTGAGAGCAACGCTGGAAGATATCATCTCCATGACGTT 14710
QY 901 TCGATTATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
14709 TCGATTATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 14650
QY 961 ACCGCGGCAACAAGAAAGCGGCTCTATGAAGATGCGTGGCGCAAGACGCTGTGAGC 1020
14649 ACCGCGGCAACAAGAAAGCGGCTCTATGAAGATGCGTGGCGCAAGACGCTGTGAGC 14590
QY 1021 AACACGATATGATGACGAGCTGGAAGCAGCATTCGTGCGAAGCAGTCACTGAGC 1080
14589 AACACGATATGATGACGAGCTGGAAGCAGCATTCGTGCGAAGCAGTCACTGAGC 14530
QY 1081 GAAGCGCTGCAATTCGCTGACGAGGTTGAAATAA 1119
14529 GAAGCGCTGCAATTCGCTGACGAGGTTGAAATAA 14491

```

RESULT 13
AL627275.C 145050 bp DNA linear BCT 06-JUN-2002
LOCUS
DEFINITION
Salmonella enterica serovar Typhi (Salmonella typhi) strain CT18,
complete chromosome, segment 11/20.
ACCESSION
AL627275 AL513382
VERSION
AL627275.1 GI:16503698
KEYWORDS
ORGANISM
Salmonella enterica subsp. enterica serovar Typhi.
Salmonella enterica subsp. enterica serovar Typhi
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Salmonella.
REFERENCE
1 (bases 1 to 145050)
Parkhill, J., Dougan, G., James, K. D., Thomson, N. R., Pickard, D.,
Wain, J., Churcher, C., Mungall, K. L., Bentley, S. D., Holden, M. T. G.,
Sebahitha, M., Baker, S., Basham, D., Brooks, K., Chillingworth, T.,
Conerton, P., Cronin, A., Davis, P., Davies, R. M., Dowd, L., White, N.,
Farrah, J., Feltwell, T., Hamlin, N., Haque, A., Hien, T. T., Holroyd, S.,
Jagels, K., Krogh, A., Larsen, T. S., Leather, S., Moule, S., O'Gaora, P.,
Parry, C., Quail, M., Rutherford, K., Simmonds, M., Skelton, J.,

TITLE Stevens, K., Whitehead, S. and Barrell, B.G.
Complete genome sequence of a multiple drug resistant Salmonella
JOURNAL enterica serovar Typhi Ctr18
MEDLINE Nature 413 (6858), 848-852 (2001)
PUBMED 21534947
REFERENCE 11677608
AUTHORS 2 (bases 1 to 145050)
TITLE Parkhill, J.
JOURNAL Direct Submission
Submitted (25-OCT-2001) Submitted on behalf of the Salmonella
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA, UK
E-mail: parkhill@sanger.ac.uk
NOTES: Details of S. typhi sequencing at the Sanger Centre are available
on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/S_typhi/).
FEATURES
source 1. 145050
/organism="Salmonella enterica subsp. enterica serovar
Typhi"
/strain="CT18"
/db_xref="taxon:90370"
636..640
/note="possible RBS"
647..1003
/gene="STY2720"
/note="yifB"
647..1003
/gene="STY2720"
/note="Orthologue of E. coli yifB (YFEB_ECOLI); Pasta hit
to YFEB_ECOLI (118 aa), 81% identity in 118 aa overlap"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAD02682.1"
/db_xref="GI:16503699"
/db_xref="SPTREMBL:O824S3"
/translation="MTLVGIRKNDITKRRNLEHGIDYRHDYRVGDIDPLPLNT
FLAEIWMQPLNRTGRTWRKLDGHRGSLTDADSAALWEPALIKRPLICAPGRFM
LIGFESRYQNFDEV"
995..1000
/gene="STY2720"
/note="possible RBS"
1007..2134
/gene="STY2721"
/note="dape"
1007..2134
/gene="STY2721"
/note="Orthologue of E. coli dape (DAPE_ECOLI); Pasta hit
to DAPE_ECOLI (375 aa), 94% identity in 375 aa overlap"
/codon_start=1
/transl_table=11
/product="succinyl-diaminopimelate desuccinylase"
/protein_id="CAD02683.1"
/db_xref="GI:16503700"
/db_xref="SPTREMBL:O824S2"
/translation="MSCPYELTQQLIRPSLSPDDVGCALMIERIKIGFTIEMD
EGDTQNVWAMRGEGTFLAFAGTDDVAVAGDWMINPPEPTIRDMGLGKGAADKQ
SLAMVAARFVAQHPHRRGLAFLITSDSEASAKNGTVKVEALMARERLDYCLV
GEPSTIEIVGVKNGRSGSLCNLTIHGVGVAVPHLADNPVHRAPLNLEVAIE
WDRGNDFEFPATSMOINAGTSGNIVGELFOPNFRTSTELTDEMKEKRVHALL
KHQRLRYDVMWISGCPFLIARGLVAVVAIHYNEIKPOLITGTGTSGRITAMG
AOYVELGPVNAITHKINECVNADQLDLNMTORIMEQLVA"
1031..1978
/gene="STY2721"
/note="Pfam match to entry PF01546 Peptidase M20.
Peptidase family M20/M25/M40, score 315.30, E-value
7.4e-91"
1187..1216
/gene="STY2721"
/note="PS00758 Arge / dape / ACY1 / CPG2 / yscs family
signature 1"

misc_feature 1295..1414
/gene="STY2721"
/note="PS00759 Arge / dape / ACY1 / CPG2 / yscs family
signature 2"
2150..2153
/note="possible RBS"
2162..2362
/gene="STY2722"
2162..2362
/gene="STY2722"
/note="Similar to Escherichia coli similar to TR_P77259
(EMBL:D90875) (66 aa) fasta scores: E(): 1e-27, 95.5% id
in 66 aa. Contains hydrophobic, possible membrane-spanning
region"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="CAD02684.1"
/db_xref="GI:16503701"
/db_xref="SPTREMBL:O824S2"
/translation="MDLAKYMWLVFLVGLVNLVYIKLRIDHKKFLANKPELP
HRDENKWDDEDMDPKKDPK"
complement(2389..4404)
/gene="STY2723"
complement(2389..4404)
/gene="STY2723"
/note="Orthologue of E. coli YPFI_ECOLI; Pasta hit to
YPFI_ECOLI (671 aa), 72% identity in 662 aa overlap"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAD02685.1"
/db_xref="GI:16503702"
/db_xref="SPTREMBL:O824S1"
/translation="MSDIDALRALTSOMQEGIRLLVTSQAWCERAEAIRAVLP
GKWWVADPAQPCRTPOALQTLIGSEPHRAIFDANQGDAAFAALSTLQASWLL
LMPYETBSRPDITSLNWDCAQPIPTPQKQKRLTSLRDPOTLLRQROPCWP
STPSRRWRPATEGPQEPQEAAILSLRLEMPGVAIVARGRKSAIAQFISOMGT
AIVTAPAKATADITLAAFGAERFCFMDALSLAGARVADLVDEAAIAPILQIVS
RFPRIILITTVQGEYEGTGFLFKFCARFQHLRFTLRQVRAPECPLENTYSEALI
FDEFAQAPDHGIIASAFYQOAMGKTPALPRAVYQLLSGAHYRTSPDLRRMDAG
OHRLOATANNRVAGALMLVEEGLSAELSAVWCGRRRRGNVQSLAHSQDLAA
TIVGRVSRIRAHPARQREGISQQLIACCEQAQCDYISVFGTTPPELIRWQKGF
VIVRWGSHREASGCTTAMALLPLSDAGRLAQOERRRLRQDIDITQNGEXIPLAA
LREQALNDEWRELWGFARFARHLTSLGLCHRLQVSAPLPALGRLEKASDAEL
CARLRISGRKALLALQRAQAAQALIDLADGRQRLRDVMPGGGDHAG"
complement(2878..3207)
/gene="STY2723"
/note="Pfam match to entry PF00583 Acetyltransf,
Acetyltransferase (GNAT) family, score 32.50, E-value
9.7e-06"
complement(3790..3813)
/gene="STY2723"
/note="PS00017 ATP/GTP-binding site motif A (P-loop)"
complement(4420..5283)
/gene="STY2724"
complement(4420..5283)
/gene="STY2724"
/note="Orthologue of E. coli P76563; Pasta hit to P76563
(287 aa), 92% identity in 286 aa overlap. Contains
hydrophobic, possible membrane-spanning region"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="CAD02686.1"
/db_xref="GI:16503703"
/db_xref="SPTREMBL:O824S0"
/translation="NRWGRRESNNVEDRRNRPQGPSLGGPCFRLPRGKGIILLVV
VVAAGYVDLTGLTGQVPSQOQSTRSISPNDDEAKFTSVLTATETWGLQFKMG
RGYQOPKLVWYRGMTGTGCGAGSVWGPYCADQTVYIDLSFVVMNKGADGDF
OGVYIAHEGVHGVOKLIGTEPRVLOQNAQSOTENRSLVREMLADCAQVWGSMD
OOGVLEAGDLEELANAAQALGDDRLQOQOQGVYVDSFTHGHSBQRYSMFKNGPFGSD
PACNTFGKNP"

```

RBS      complement(5287..5292)
          /note="possible RBS"
          /complement(5415..6128)
          /gene="STY2725"
          /note="STY2725"
          /note="purc"
          /complement(5415..6128)
          /gene="STY2725"
          /note="Orthologue of E. coli purc (PUR7_ECOLI); fasta hit
          to PUR7_ECOLI (237 aa), 95% identity in 237 aa overlap"
          /codon_start=1
          /transl_table=11
          /product="phosphoribosylaminoimidazole-succinocarboxamide
          synthase"
          /protein_id="CAD02687.1"
          /db_xref="GI:16503704"
          /db_xref="SPTREMBL:O8Xf69"
          /translation="MOKAELKRGKAKVYSENPDLVLEPRNDSAGDARIEOPD
          RGVNANKFNHIMTKLAAGLPTOMERLSDTELVKLEKVPYECVRRNRAQSLV
          KRLVEEGMELNPLFDELFLKNDALHDMVNSVCEYTGWSQENLAKMKELTYKAND
          VLKLFDDAGLILVDFKLEFGLYKGEVVLGDEFSPGSRIMDKETLMDKDRFQSL
          GGLIAYAVAVARHLGKLD"
          /complement(5421..6122)
          /gene="STY2725"
          /note="Pfam match to entry PF01259 SAICAR synt, SAICAR
          synthetase, score 414.20, E-value 1.2e-120"
          /complement(5586..5612)
          /gene="STY2725"
          /note="PS01058 SAICAR synthetase signature 2"
          /complement(5829..5873)
          /gene="STY2725"
          /note="PS01057 SAICAR synthetase signature 1"
          /complement(6137..6141)
          /note="possible RBS"
          /complement(6303..7337)
          /gene="STY2726"
          /note="nlp"

Query Match      79.1%; Score 885.4; DB 1; Length 145050;
Best local Similarity 87.0%; Pred. No. 7.8e-212;
Matches 973; Conservative 0; Mismatches 146; Indels 0; Gaps 0:

QY      1 ATGCATACCAGGCTCCATTCATCAAGTAGAATAACAACGATTTACGTGGGAATG 60
          |||||||
Db 64624 ATGCATACCAGGCTCCATTCATCAAGTAGAATAACAACGATTTACGTGGGAATG 64565

QY      121 GTGGAAGCAACGGTCAATCAATCAAGGCGGTGAACGCGCTTGCGTATATCGTCCGT 180
          |||||||
Db 64504 GTGGAAGCAACGGTCAATCAATCAAGGCGGTGAACGCGCTTGCGTATATCGTCCGT 64445

QY      181 GTATCCGTCACCGAGATGACGCGCGCAAGAGCGTTCAAACTCATCAACGCAAGGTTAAC 240
          |||||||
Db 64444 GTTTCGTACCGACATGACGCGCGCAAGAGCGTTCAAACTCATCAACGCAAGGTTAAC 64385

QY      241 GTGCGCGTGTGGTGATCCATCCATTCGATTCGATTCGCTGCGTGAAGAAGCGAATAC 300
          |||||||
Db 64384 GTGCGCGTGTGGTGATCCATTCGATTCGATTCGCTGCGTGAAGAAGCGAATAC 64325

QY      301 GGGCGTCATTGTCTGCGTATTAACCTGCAATATCGGTATGAAGAGCGTATTCGCAIG 360
          |||||||
Db 64324 GGGCGTCATTGTCTGCGTATTAACCTGCAATATCGGTATGAAGAGCGTATTCGCAIG 64265

QY      361 GTGCTTACCTGTGCGCGCGCAATAACATTCGATTCGCTTATGGCGTTAAGCGCGATTCG 420
          |||||||
Db 64264 GTGCTTACCTGTGCGCGCGCAATAACATTCGATTCGCTTATGGCGTTAAGCGCGATTCG 64205

QY      421 CTGGAATAAAGATCTGCAAGAAGATATGAGGAACCGACGCGAGCGGTGCTGTAATCT 480
          |||||||
Db 64204 CTGGAATAAAGATCTGCAAGAAGATATGAGGAACCGACGCGAGCGGTGCTGTAATCT 64145

```

```

QY      481 GCCATGCGTATCTGTTGATCATCTGATCGCCTGAACCTTCGATCACTTCAAGTCAGCGTG 540
          |||||||
Db 64144 GCCATGCGCGCATGTCGATCATCTGATCGCCTGAACCTTCGATCACTTCAAGTCAGCGTG 64085

QY      541 AAGCGCTGACGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
          |||||||
Db 64084 AAGCGCTGACGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 64025

QY      601 CAGCGCTTCATCTGCGGATACACCAAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
          |||||||
Db 64024 CAGCGCTTCATCTGCGGATACACCAAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 63965

QY      661 GCCATTGTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
          |||||||
Db 63964 GCCATTGTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 63905

QY      721 GGGCGCGACCGCTGGAAGAGATCAAGCGCTTCGATTAATTTGAATTCGCTGCTGCTGCTGCTG 780
          |||||||
Db 63904 GGGCGCGACCGCTGGAAGAGATCAAGCGCTTCGATTAATTTGAATTCGCTGCTGCTGCTGCTGCTG 63845

QY      781 CGTTCGCGAGGATCACTTCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
          |||||||
Db 63844 CGTTCGCGAGGATCACTTCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 63785

QY      841 ATCGCTACGTTTACGCGCTGCGAGCAACGCTTCGGAAGATATCATCTCCATTCGATTCGATTCG 900
          |||||||
Db 63784 ATCGCTACGTTTACGCGCTGCGAGCAACGCTTCGGAAGATATCATCTCCATTCGATTCGATTCG 63725

QY      901 TCGATTATCGCGTGGTGTGTAATGCGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
          |||||||
Db 63724 TCGATTATCGCGTGGTGTGTAATGCGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 63665

QY      961 ACCGCGCGCAACAGAAAGCGCGCTCTATGAAGATGCGCGCAACGCGCTGCTGAGC 1020
          |||||||
Db 63664 ACCGCGCGCAACAGAAAGCGCGCTCTATGAAGATGCGCGCAACGCGCTGCTGAGCAGACGATCGAT 63605

QY      1021 AACACGATATGATGACACGACGCTGGAAGCAACGCTTCGCGAAGACCGATGCTGAGC 1080
          |||||||
Db 63604 AACGACGATATGATGACGCGCTGGAAGCAACGCTTCGCGAAGACCGATGATGCTGATGAT 63545

QY      1081 GAAGCGCGTGAATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 1140
          |||||||
Db 63544 GAAGCGCGTGAATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 63506

RESULT 14
LOCUS      D90881/c
DEFINITION E.coli genomic DNA, kohara clone #428(56.8-57.0 mln.).
ACCESSION D90881 AB001340
VERSION    D90881.1 GI:1799913
KEYWORDS   Complete and shotgun sequencing; PBPE; POMA; gppe; ndk; yfgA; yfgB.
SOURCE     Escherichia coli (strain:K12) DNA, clone_11b:kohara lambda miniset
           library clone:kohara clone #428.
ORGANISM   Escherichia coli
           Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
           Escherichia.
REFERENCE  1 (sites)
AUTHORS    Yamamoto,Y., Alba,H., Baba,T., Hayashi,K., Inada,T., Isono,K.,
           Itoh,T., Kimura,S., Kitagawa,M., Makino,K., Miki,T., Mitsuhashi,N.,
           Mizobuchi,K., Mori,H., Nakade,S., Nakamura,Y., Nishimoto,H.,
           Oshima,T., Oyama,S., Saito,N., Sempel,G., Setoh,Y., Takemoto,K.,
           Sivasundaram,S., Tagami,H., Takahashi,H., Takeda,Y., Takemoto,K.,
           Uehara,K., Wada,C., Yamagata,S., and Horiiuchi,T.
           Construction of a contiguous 874-kb sequence of the Escherichia
           coli -K12 genome corresponding to 50.0-68.8 min on the linkage map
           and analysis of its sequence features
           DNA Res. 4 (2), 91-113 (1997)
JOURNAL    MEDLINE
REFERENCE  2 (sites)
AUTHORS    Alba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A., Horiiuchi,T.,
           Ikemoto,K., Inada,T., Isono,K., Isono,S., Itoh,T., Kanai,K.,
           Kasai,H., Kashimoto,K., Kim,S., Kimura,S., Kitagawa,M.,

```

Kitakawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H.,
 Motomura, K., Nakamura, Y., Nishimoto, Y., Nishio, Y., Oshima, T.,
 Saito, N., Sampei, G., Seki, Y., Tagami, H., Takemoto, K., Wada, C.,
 Yamamoto, Y. and Yano, M.
 The systematic sequencing of the *Escherichia coli* genome in Japan
 unpublished
 3 (bases 1 to 4775)
 Mori, H.
 Direct Submission
 Submitted (23-JAN-1997) Hirotsada Mori, NARA Institute of Science
 and Technology, Res. & Edu. Center for Genetic Info., 8916-5
 Takayama, Ikoma, Nara 630-01, Japan
 (E-mail:hmori@gtc.aist-nara.ac.jp, Tel:81-7437-2-5660,
 Fax:81-7437-2-5669)
 Collaboration information:
 Project:
 The Japan E.coli genome DNA sequencing project
 Group:
 The Japan E.coli genome DNA sequencing group
 Members: (1995.4 - 1996.3)
 Alba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A.,
 Horiiuchi, T., Ikemoto, K., Inada, T., Isono, K., Isono, S.,
 Itoh, T., Kanai, K., Kasai, H., Kashimoto, K., Kim, S.,
 Kimura, S., Kitagawa, M., Kitakawa, M., Makino, K.,
 Masuda, S., Miki, T., Mizobuchi, K., Mori, H., Motomura, K.,
 Nakamura, Y., Nishimoto, H., Nishio, Y., Oshima, T., Saito, N.,
 Sampei, G., Seki, Y., Tagami, H., Takemoto, K., Wada, C.,
 Yamamoto, Y. and Yano, M.
 Headed by:
 Name: Takashi Horiiuchi
 Address: National Institute of Basic Biology, Okazaki, 444, Japan
 E-mail: kishori@nibb.ac.jp
 Information operator:
 Name: Hirotsada Mori
 Address: NARA Institute of Science and Technology,
 Ikoma, 630-01, Japan
 E-mail: hmori@gtc.aist-nara.ac.jp
 URL:
 The Japan E. coli genome database
 http://b3w3.aist-nara.ac.jp.
 Location/Qualifiers
 1..4775
 /organism="Escherichia coli"
 /strain="K12"
 /db_xref="taxon:562"
 /map="56.8 min"
 /clone="Kohara clone #428"
 /clone_lib="Kohara lambda miniset library"
 /note="Nucleotide position 2635874-2640648 from the
 initiation site of Thra (0 min.). This clone is from
 Kohara lambda miniset library"
 complement(1..710)
 /gene="gcpe"
 /gene="gcpe"
 complement(<1..710)
 /gene="gcpe"
 /note="similar to [PIR Accession Number S23058]"
 /codon_start=1
 /transl_table=11
 /product="gcpe protein"
 /protein_id="BA20919.1"
 /db_xref="GI:4433789"
 /translation="MHNQAPIQRRKSTRIVYGVPGCAPTAVOSMTNTRTTDEAT
 VNOIKALERVADIVRVSVPTMDAAEFLLIKOOVNPVLAVIDHEPYRALVAEYGV
 DCLIRNPGNIGNEIRIMVDCARDKNIRIGVNGSLEKDIQEKYGEPTQALLES
 AARVHDRLRNFDOFKVSKASDVFLAVESRYLLAKQIDQPLHICITBAGARSGAV
 KSAIGLISGIGDILR"
 complement(737..1750)
 /gene="yfga"
 /gene="yfga"
 complement(737..1750)
 /gene="yfga"
 complement(737..1750)
 /note="similar to [SwissProt Accession Number P27434]"
 /codon_start=1
 /transl_table=11

/protein_id="BA16403.1"
 /db_xref="GI:1799915"
 /translation="MNEARHONENALTGARLRNAREQIGLSQAAERELCKNVTY
 RDIIEBKAPADIASTFIRKTIYRARIYIPHEELLPGLCKQAPLRARAKVAPQSSLS
 GKRRKKGDMGLMTFTWLVLVAVIGLSGAMWQRRKQOEITITMAQSSAELSSNEQ
 GQSVPLNTSTTDPATSTPPASVDATATWOTPAVTAAPADVPQONAVSPSQANV
 DTATAPATPAATPDGAAPLPTDQAGVTTVPADPNALVMNFTADMLLEVDTAGKLEF
 SGMRKQGNLNTGQAPYKLTICAPAAVQIYQYGGKVPDLSRFIRTNQVARTLTMARQS
 PAQ"
 complement(2035..3189)
 /gene="yfgb"
 complement(2035..3189)
 /gene="yfgb"
 /note="similar to [SwissProt Accession Number P36979]"
 /codon_start=1
 /transl_table=11
 /protein_id="BA16404.1"
 /db_xref="GI:1799916"
 /translation="MSEQLVTPENVTKDKINLIDLNRQMRFEFFKDLGKPFRAQ
 VMKMYHYCCDNFDEMIDINKVLKLEVAETRAVEVEEORSQGTIKMALVAGDO
 RVEITYIPEDDRATLCVSSQVGCALCKFCSTAOGFNNRLRYSEIIGVVRKATYV
 AAKVTGQRPRTTNVMMGEPILNINNVNPAEMIMDDCGFLSKRRVTLSTQGYPA
 LDKLGIMIDVALAISLHAPDELRDITVINKTYNETFLAARVRLKESNMQGKVT
 IETVMLDHNVDGTEHQAELKIDTPCKINILPMVPEFGATYGRSSNSRIDFSKVL
 MSYGFITVAKRTKGDIDAACGQAGDVIDRTKRRLRKRGQALDIAKAV"
 complement(3339..3770)
 /gene="ndk"
 complement(3339..3770)
 /gene="ndk"
 /note="similar to [PIR Accession Number JH0495]"
 /codon_start=1
 /transl_table=11
 /product="nucleoside-diphosphate kinase (EC 2.7.4.6)"
 /protein_id="BA16405.1"
 /db_xref="GI:1799917"
 /translation="MAIERTFSLIKNAVAKVNTGIRPFEAGKRYTGTMHLITV
 EQARKFYAHDKPFPDGLVEFTSGPIVSVLEGNAAVQRHDLGATVPANALAGT
 LRADYADSLTEGTGHSDESASARIAYFEGECPRTR"
 complement(3919..4775)
 /gene="pppf"
 /note="PBPf"
 complement(3919..4775)
 /gene="pppf"
 /note="PBPf"
 complement(3919..4775)
 /note="similar to [SwissProt Accession Number P38050]"
 /codon_start=3
 /transl_table=11
 /product="PENICILLIN-BINDING PROTEIN 1A (PBP-1A)."
 /protein_id="BA16406.1"
 /db_xref="GI:1799918"
 /translation="RRIMADEAQLPDSALPRVAPLWMTGTSYGRDAMAGVNAVY
 VIGIWTGRPDGPVVGQFGFASAVPLNQNVLILSRANLFEDEPRNSVTRGVICMP
 GGOSLPEEGDNCRRRLATWLIDGSOPTLLEQGINIRPIMLDENGRKVAADCP
 GARQEMINWPLPLEPMLPASRRRAVRPRPASTSCPYGHDAQPLQLTGVYDGLTIK
 RLPQAAEATLPIQSSGSGAGERWFLNGEPLTERGNVTLHLTDKDYQLVMDVDGQI
 ATVKRFVMO"
 BASE COUNT 1092 a 1304 c 1228 g 1151 t
 ORIGIN
 Query Match 63.4%; Score 710; DB 1; Length 4775;
 Best Local Similarity 100.0%; Pred No. 8.8e-166;
 Matches 710; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 ATGCATACAGCGCTCCCAATTCACGTAGAAATCAACAGCTATTACGTGGGAATGTG 60
 Db 710 ATGCATACAGCGCTCCCAATTCACGTAGAAATCAACAGCTATTACGTGGGAATGTG 651
 QY 61 CCGATTGGCGATGCTGCCATCGCGTACAGTCCATACCAATACGCGTACGACAGAC 120
 Db 650 CCGATTGGCGATGCTGCCATCGCGTACAGTCCATACCAATACGCGTACGACAGAC 591
 QY 121 GTGCAAGCAAGCGTCAATCAATCAAGGCGGTGCAAGCGGTGGCGGTGATVTCGTCGCT 180
 Db 590 GTGCAAGCAAGCGTCAATCAATCAAGGCGGTGCAAGCGGTGGCGGTGATVTCGTCGCT 531

Mon May 12 08:44:41 2003

us-09-921-992-3.rge

Page 21

Search completed: May 10, 2003, 16:14:16
Job time : 3624 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2003, 14:31:07 ; Search time 320 Seconds

(without alignments)
7874.957 Million cell updates/sec

Title: US-09-921-992-3

Perfect score: 1119

Sequence: 1 atgcataaccgctccat.....ttcagcaggttgaataaa 1119

Scoring table: IDENTITY-NUC

Gapop 10.0 , Gapept 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq_101002:*

1: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1119	100.0	1119	21	AAA95478
2	1119	100.0	1119	21	AAAB8705
3	1119	100.0	1119	22	AA82653
4	1119	100.0	1119	24	AA831203
5	622	55.6	1830121	17	AAT42063
6	411.8	36.8	5484	23	AA573006
7	403.6	36.1	640681	24	ABA92787
8	382.6	34.2	33140	22	AA28536
9	356	31.8	1857	23	AA590066

10	293.4	26.2	1083	24	ABK74401
11	254.6	22.8	6157	24	ABO70939
12	251.4	22.5	1107	24	ABO67966
13	251.4	22.5	1107	24	ABO69993
14	244	21.8	35829	23	AA555573
15	243.8	21.8	1164	22	AAH52056
16	243.6	21.8	994	24	ABQ68324
17	239.4	21.4	1134	22	AAH67170
18	239.4	21.4	349980	22	AAH68530
19	239.4	21.4	349980	22	AAH68531
20	211.8	18.9	1180	19	AAK14017
21	136.4	12.2	507	21	AACT5580
22	136.4	12.2	507	24	ABN76496
23	88	7.9	1152	23	AA588480
24	87.2	7.8	2535	20	AAK20674
25	72.4	6.5	4467	10	AAK92428
26	70.6	6.3	584	24	AAK31220
27	68.4	6.1	596	24	AAK31222
28	68.4	6.1	670	24	AAK31221
29	58.8	5.3	60873	21	AAK81469
30	58.8	5.3	349980	21	AAK21610
31	56.8	5.1	348	24	ABK78811
32	53.4	4.8	1038602	20	AAZ01425
33	52.6	4.7	293	24	AAK31224
34	52.6	4.7	456	24	AAK31225
35	52.6	4.7	504	16	AAQ99779
36	50.8	4.5	1398	20	AAK20609
37	50	4.5	403	24	AAK31223
38	48.8	4.4	564	24	AAK31234
39	48.4	4.3	379	24	AAK31215
40	48.4	4.3	528	24	AAK31214
41	48.4	4.3	2520	24	AAK31201
42	48.2	4.3	353	24	AAK31238
43	48.2	4.3	464	24	AAK31232
44	47.2	4.2	601	24	AAK31210
45	45.4	4.1	613	24	AAK31231

ALIGNMENTS

RESULT 1
AAA95478
ID AAA95478 standard; DNA; 1119 BP.
XX
AC AAA95478;
XX
DT 27-FEB-2001 (first entry)
XX
DE E. coli essential gene gcpe.
XX
KW Bacterial growth; inhibitor; ygbB; yfnc; yacB; ychB; yejD; yrfI;
KW yggT; yjeF; yjaO; yrcG; yhbC; ygbP; yjeY; gcpe; kdsB; pfs; yaaI;
KW b180a; yeaA; yagF; b1983; yldD; yocG; yjbc; antibacterial compound;
KW H. influenza; S. pneumoniae; H. pylori; B. burgdorferi; pesticide;
KW M. tuberculosis; antibiotic; ds.
XX
OS Escherichia coli.
XX
PN WO200061793-A2.
XX
PD 19-OCT-2000.
XX
PF 07-APR-2000; 2000MO-EP03135.
XX
PR 09-APR-1999; 99EP-0107031.
XX
PR 04-FEB-2000; 2000EP-0102111.
XX
PA (GPCB-) GPC BIOTECH AG.
XX
PI Loferer H, Jacobl A;
XX
DR WPI; 2000-687048/67.

Bacillus lichenif
Listeria monocyt
Listeria monocyt
Listeria monocyt
Propionibacteri
Mycobacterium tub
Listeria monocyt
C glutamicum codin
C glutamicum codin
H. pylori GHO 76
Human ORF143 CDNA
DNA encoding novel
Polynucleotide seq
Sequence encoding
Zea mays partial g
Zea mays partial g
Zea mays partial g
N. meningitidis pa
Neisseria meningit
Bacillus clausii g
Complete genome se
Zea mays partial g
Zea mays partial g
Cross-reactive all
Polynucleotide seq
Zea mays partial g
Glycine max partia
Arabidopsis thalia
Arabidopsis thalia
Glycine max partia
Pinus taeda partia
Zea mays partial g

Identifying antibacterial compounds, comprises identifying an antagonist or inhibitor of the expression of a gene encoding a polypeptide essential for bacterial growth or survival.

Claim 1: Fig 1; 75bp; English.

The present invention relates to antagonists and inhibitors of 24 bacterial genes and proteins. The proteins are thought to be essential for growth in several species of bacteria (including *S. pneumoniae*, *B. burgdorferi*, *H. influenza* and *H. pylori*). The proteins and coding sequences shown in the specification can be used to identify antagonists and inhibitors which can be used in disease treatment and pesticides. In particular, they can be used against *M. tuberculosis*. The present sequence is one of the genes of the invention.

Sequence 1119 BP; 272 A; 279 C; 317 G; 251 T; 0 other;

Query Match 100.0%; Score 1119; DB 21; Length 1119;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0

1 ATGCATTAACACAGGCTTCATATTCACGTACAAATATCAACACGTAATTACGTTGGGAATG 60
1 ATGCATTAACACAGGCTTCATATTCACGTACAAATATCAACACGTAATTACGTTGGGAATG 60

61 CCGATTGGCGATGAGTGTGCTGGCATCCGCCGTACAGTCCATACCAATATCGGCTACGACAGC 120
61 CCGATTGGCGATGAGTGTGCTGGCATCCGCCGTACAGTCCATACCAATATCGGCTACGACAGC 120

121 GTGCAAGCAGCGGTCAATATTAATCAAGCGCTGGAACGCGTGGCGCTGATATCGTCGT 180
121 GTGCAAGCAGCGGTCAATATTAATCAAGCGCTGGAACGCGTGGCGCTGATATCGTCGT 180

181 GTATCCGTATCCAGCATGTGAGAGCGGAGCGGACAGAGGTTCAAACTCATCAACACAGAGTTAC 240
181 GTATCCGTATCCAGCATGTGAGAGCGGAGCGGACAGAGGTTCAAACTCATCAACACAGAGTTAC 240

241 GTGCCGTGTGTGGCTGACATTCGACTTTCGACTATTCGATTCGCGCTGAAGATAGCGGATAC 300
241 GTGCCGTGTGTGGCTGACATTCGACTTTCGACTATTCGATTCGCGCTGAAGATAGCGGATAC 300

301 GCGCGTCGATGTCGTGCGTATTAACCTCGGCATATTCGGAATGAAGAGCGTATTCGCATG 360
301 GCGCGTCGATGTCGTGCGTATTAACCTCGGCATATTCGGAATGAAGAGCGTATTCGCATG 360

361 GTGCTTGAAGTGTGCGCGGATTAATAACATATTCGATTCGATTCGCGCTGAAGATAGCGGATAC 420
361 GTGCTTGAAGTGTGCGCGGATTAATAACATATTCGATTCGATTCGCGCTGAAGATAGCGGATAC 420

421 CTGGAATAAGATCTCTGCAAGAAAGTATGCGCAACCGACGCGCTGCTGGAATCT 480
421 CTGGAATAAGATCTCTGCAAGAAAGTATGCGCAACCGACGCGCTGCTGGAATCT 480

481 GCGATGCGTCAATGTTGATCACTCATGATGCGCTGCGCTGATCACTTCAAAATCAGCGTG 540
481 GCGATGCGTCAATGTTGATCACTCATGATGCGCTGCGCTGATCACTTCAAAATCAGCGTG 540

541 AAGCGTCTGACGCTTCCTGCGCTGTGATGCTTATTCGTTGCTGGCAAAACAGATCAT 600
541 AAGCGTCTGACGCTTCCTGCGCTGTGATGCTTATTCGTTGCTGGCAAAACAGATCAT 600

601 CAGCGCTTGCATCTGCGGAGTACCGAAGCGGCTGCTGGCGGACGCGGAGGAGTAAATCC 660
601 CAGCGCTTGCATCTGCGGAGTACCGAAGCGGCTGCTGGCGGACGCGGAGGAGTAAATCC 660

661 GCGATTTGTTAGTCTGCTGCTGTCTTAAGGCAATCGGCGACAGCCTCGCGTATCGCTG 720
661 GCGATTTGTTAGTCTGCTGCTGTCTTAAGGCAATCGGCGACAGCCTCGCGTATCGCTG 720

721 GCGGCGGATCCGGTCTGGAAGGATCAAGTCCGTTTCGATATTTTGAATTCGCTGCGTATC 780
721 GCGGCGGATCCGGTCTGGAAGGATCAAGTCCGTTTCGATATTTTGAATTCGCTGCGTATC 780

QY	781	CGTTGCGCGGAGTCAACTTCATCGCCCGCCGACCTGTGCGTCAAGAAATTGGATT	840
Db	781	CGTTGCGCGGAGTCAACTTCATCGCCCGCCGACCTGTGCGTCAAGAAATTGGATT	840
QY	841	ATCGGTACGGTTAACCGCCTGGAGAACCGCCTGGAAATATCATCTCGATGACGTT	900
Db	841	ATCGGTACGGTTAACCGCCTGGAGAACCGCCTGGAAATATCATCTCGATGACGTT	900
QY	901	TGCGATTATCGCGTGGTGGTGAATGCGCCAGGTAGGCGCGTGGTTTCTACACTGGCGTC	960
Db	901	TGCGATTATCGCGTGGTGGTGAATGCGCCAGGTAGGCGCGTGGTTTCTACACTGGCGTC	960
QY	961	ACCGGCGGCAACAAAGAAAGCGGCGCTCTATGAAATGAGCGCGCAAGAACCGCTGAC	1020
Db	961	ACCGGCGGCAACAAAGAAAGCGGCGCTCTATGAAATGAGCGCGCGCAAGAACCGCTGAC	1020
QY	1021	AACAACGATATGATGACGACGACTGGAAAGCAGCATTTGTGTGGAAAGCAGTCAGTGGAC	1080
Db	1021	AACAACGATATGATGACGACGACTGGAAAGCAGCATTTGTGTGGAAAGCAGTCAGTGGAC	1080
QY	1081	GAGCGCGCTCGAATTGACGTTGACGAGTGTGAAAAATTA 1119	
Db	1081	GAGCGCGCTCGAATTGACGTTGACGAGTGTGAAAAATTA 1119	
RESULT 2			
ID	AAA88705	standard; DNA; 1119 BP.	
AC	AAA88705;		
DT	05-FEB-2001	(first entry)	
DE	E. coli FUN essential gene gcpe.		
KW	FUN gene; gcpe gene; essential gene; antibacterial; antibiotic;		
KW	screening; infection; therapy; antagonist; surrogate marker; ds.		
OS	Escherichia coli.		
PN	EPI043403-A1.		
PD	11-OCT-2000.		
PE	09-APR-1999;	99EP-0107031.	
PR	09-APR-1999;	99EP-0107031.	
PA	(GPCG-) GPC GENOME PHARM CORP AG.		
PI	Not given;		
XX	WPI: 2000-640125/62.		
XX	The present sequence is that of the Escherichia coli FUN gene		
XX	gcpe, which encodes a protein that is essential for bacterial		
XX	growth or survival. gcpe is 1 of 22 E. coli genes (see		
XX	AA88692-713) identified as being essential (there is no deletion		
XX	genotype). These 22 genes fulfil criteria for being attractive		
XX	antibacterial targets: hypothetical open reading frames coding for		
XX	essential functions (mutation is lethal for growth in rich media);		
XX	broad conservation (orthologues are present in a wide range of		
XX	bacteria, including Haemophilus influenzae, Streptococcus pneumoniae,		
XX	Helicobacter pylori and Borrelia burgdorferi); and low toxicity		
XX	potential in higher organisms (mostly no orthologues were identified		
XX	test molecule		
XX	Claim 1; Page 25; 55pp: English.		

CC In *Saccharomyces cerevisiae*). An antagonist or inhibitor of the
CC expression of an essential gene or of its function provides the key
CC for antibacterial therapy. The invention provides methods for
CC identifying such antagonists or inhibitors. These involve
CC contacting a bacterial cell comprising an essential gene with a
CC candidate antagonist or inhibitor, and testing whether contact leads
CC to cell growth inhibition and/or cell death. The method allows the
CC development of new broad spectrum antibiotics. A conditional mutant
CC of an essential gene can be used to induce a lethal phenotype in
CC bacteria for the analysis of surrogate markers.

XX Sequence 1119 BP: 272 A; 279 C; 317 G; 251 T; 0 other;

Query Match 100.0%; Score 1119; DB 21; Length 1119;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCATACCAAGGCTCCATTCACGTGAAATCAACACGATATTACGTTGGAGATG 60
Db 1 ATGCATACCAAGGCTCCATTCACGTGAAATCAACACGATATTACGTTGGAGATG 60
Qy 61 CCGATTGGCGATGTCCTCCATTCACGTGAAATCAACACGATATTACGTTGGAGATG 120
Db 61 CCGATTGGCGATGTCCTCCATTCACGTGAAATCAACACGATATTACGTTGGAGATG 120
Qy 121 GTGCAAGCAAGCGTCAATCAATCAAGCGCTGAGACGCTGATATCGTCCT 180
Db 121 GTGCAAGCAAGCGTCAATCAATCAAGCGCTGAGACGCTGATATCGTCCT 180
Qy 181 GTATCCGATACGACGATGACGCGGAGAGCGTTCAACATCATCAACAGCGATTAC 240
Db 181 GTATCCGATACGACGATGACGCGGAGAGCGTTCAACATCATCAACAGCGATTAC 240
Qy 241 GTGCGCGCTGTCGTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 300
Db 241 GTGCGCGCTGTCGTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 300
Qy 301 GCGCTGATTCCTGCTGATTAACCTGACATATCGTAATGAGAGCGTATTGCGATG 360
Db 301 GCGCTGATTCCTGCTGATTAACCTGACATATCGTAATGAGAGCGTATTGCGATG 360
Qy 361 GTGCTGATTCCTGCTGATTAACCTGACATATCGTAATGAGAGCGTATTGCGATG 420
Db 361 GTGCTGATTCCTGCTGATTAACCTGACATATCGTAATGAGAGCGTATTGCGATG 420
Qy 421 CTGGAAGAAAGATCTCAAGAAAGATGAGGAGACGCGCGCGCTGCTGATCT 480
Db 421 CTGGAAGAAAGATCTCAAGAAAGATGAGGAGACGCGCGCGCTGCTGATCT 480
Qy 481 GCGATGCGTCATGTTGATCATCTCGATCGCTGAACTTGATGATCAAGTCAAGCTG 540
Db 481 GCGATGCGTCATGTTGATCATCTCGATCGCTGAACTTGATGATCAAGTCAAGCTG 540
Qy 541 AAAGCGCTGACGCT 600
Db 541 AAAGCGCTGACGCT 600
Qy 601 CAGCGCTGATCTGGGAGTCAACGAGCGGAGGAGCGGAGCGGAGCGGAGCGGAGCGG 660
Db 601 CAGCGCTGATCTGGGAGTCAACGAGCGGAGGAGCGGAGCGGAGCGGAGCGGAGCGG 660
Qy 661 GCGATGCTGTTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 661 GCGATGCTGTTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Qy 721 GCGGCGGATCCGATCGAAGATCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 GCGGCGGATCCGATCGAAGATCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 781 CGTTGCGGAGGATCAACATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 840
Db 781 CGTTGCGGAGGATCAACATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 840

Qy 841 ATCGGTACGTTTAACGCGCTGAGACACGCTGGAAGATATCATCTCCGATGAGCTT 900
Db 841 ATCGGTACGTTTAACGCGCTGAGACACGCTGGAAGATATCATCTCCGATGAGCTT 900
Qy 901 TCGATTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 TCGATTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Qy 961 ACCGCGGCAACAGAAAGAGCGCTCTATGAGATGCGCGCAAGACCGTCTGAGC 1020
Db 961 ACCGCGGCAACAGAAAGAGCGCTCTATGAGATGCGCGCAAGACCGTCTGAGC 1020
Qy 1021 AACACGATATGATGACACGCTGAGACGATTCGTCGAAACGATGACGCTGAC 1080
Db 1021 AACACGATATGATGACACGCTGAGACGATTCGTCGAAACGATGACGCTGAC 1080
Qy 1081 GAAGCGCTGCAATTGACGCTGACGAGTTGAAAAATTA 1119
Db 1081 GAAGCGCTGCAATTGACGCTGACGAGTTGAAAAATTA 1119

RESULT 3

AAC82653
ID AAC82653 standard; DNA; 1119 BP.

XX AAC82653;

XX 15-MAR-2001 (first entry)

DE E. coli gcpE DNA.

XX Isoprenoid; gcpE; yfgB; antimicrobial; transgenic plant; agriculture;
XX antimycotic; antiparasitic; antiviral; fungicidal; herbicidal; ds.

OS *Escherichia coli*.

PN WO200072022-A1.

PD 30-NOV-2000.

PF 20-MAY-2000; 2000WO-EP04592.

XX 21-MAY-1999; 99DE-1023567.

PR 21-MAY-1999; 99DE-1023568.

PA (JOMAA/) JOMAA H.

PI Jomaa H;

DR WPI; 2001-025196/03.

DR P-PSDB; AAB45692.

PT Incorporating gcpE and yfgB genes into viruses and cells, for
PT increasing isoprenoid content and identifying e.g. antimicrobial
PT agents, comprises using DNA sequences from bacteria or parasites

PS Claim 3; Page 13-15; 36pp; German.

This invention describes a novel method for incorporating gcpE and yfgB
genes into viruses and cells for increasing isoprenoid content and
identifying e.g. antimicrobial agents, comprising using DNA sequences (I)
from the gcpE or yfgB genes of bacteria or parasites or DNA sequences
(II) which hybridize to the specified genes or encode a plasmid protein
with the same biological activity as those encoded by the genes. The
invention also describes (1) plant cells containing (I) or (II); (2)
transformed plant cells, and transgenic plants regenerated from them,
that contain (I) or (II); (3) determining the enzymatic activity of a
gcpE protein; or (4) screening compounds (A) that have antimycotic,
antiparasitic or antiviral activity in humans or animals or antiviral,
antiparasitic, fungicidal or herbicidal activity in plants. (I) and (II)
are used: (i) to increase the isoprenoid levels in viruses and cells;
(ii) for determining the enzymatic activity of gcpE and yfgB proteins;
and (iii) to identify compounds that inhibit activity of gcpE, i.e.

CC potential antibacterial, antimycotic, antiparasitic or antiviral agents
CC for use in humans or animals, or antiviral, antiparasitic, fungicidal
CC or herbicidal agents for agriculture.

SQ Sequence 1119 BP; 272 A; 279 C; 317 G; 251 T; 0 other;

Query Match	100.0%;	Score 1119;	DB 22;	Length 1119;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1119;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

OY	1	ATGCATTAACCGAGGCTCCCAATTCACGCTGTGAATAATCAACAGTATTACGTTGGGAATGTG	60
Dp	1	ATGCATTAACCGAGGCTCCCAATTCACGCTGTGAATAATCAACAGTATTACGTTGGGAATGTG	60
OY	61	CCGATTGGCGATGTGTCTCCATCCGCGGTACAGTCCATACGACCAATACGCCATTACGACAGAC	120
Dp	61	CCGATTGGCGATGTGTCTCCATCCGCGGTACAGTCCATACGACCAATACGCCATTACGACAGAC	120
OY	121	GTCGAAGCAAGCGTTCATCAATCAAGGCGTGGAGCGCTGGAGCGCTGATATGTCCTGT	180
Dp	121	GTCGAAGCAAGCGTTCATCAATCAAGGCGTGGAGCGCGTGGAGCGCTGATATGTCCTGT	180
OY	181	GTATCCGTACCGACGATGGAGCGGCGAGAACCGTTTCAAACTCATCAAAACACACAGTTTAC	240
Dp	181	GTATCCGTACCGACGATGGAGCGGCGAGAACCGTTTCAAACTCATCAAAACACACAGTTTAC	240
OY	241	GTGCCGCTGTGGCTGACATCCACTTCGATCTATGCGATTGGCGTGAAAGTAGAGCGGAATAC	300
Dp	241	GTGCCGCTGTGGCTGACATCCACTTCGATCTATGCGATTGGCGTGAAAGTAGAGCGGAATAC	300
OY	301	GAGCGCATTTCTCAGCTATTAAACCTGGGCAATTCGGTATATGAAGACCGTATTCGCATG	360
Dp	301	GAGCGCATTTCTCAGCTATTAAACCTGGGCAATTCGGTATATGAAGACCGGATTCGCATG	360
OY	361	GTGGTTGACTGTGGCGCGGATTAATAAAGATTCCGATTCGGTATGGCGTTAAACGCCGAGATCG	420
Dp	361	GTGGTTGACTGTGGCGCGGATTAATAAAGATTCCGATTCGGTATGGCGTTAAACGCCGAGATCG	420
OY	421	CTGGAAAAAGATCTGCAAGAAAAATATGGCGAACCGCGCGAGCGCGCTGTCTGGAATCT	480
Dp	421	CTGGAAAAAGATCTGCAAGAAAAATATGGCGAACCGCGCGAGCGCGCTGTCTGGAATCT	480
OY	481	GCCATTCGCTCATGTGTGATCATCTGATACGCTGAACTTCGATTCGATTCGATTCGATTCGATTC	540
Dp	481	GCCATTCGCTCATGTGTGATCATCTGATACGCTGAACTTCGATTCGATTCGATTCGATTCGATTC	540
OY	541	AAAGCGTCTGACGCTTCCTCCTCGCTGTGAGTCTTATGCTTGGCTGGCGAAAAACAGATCGAT	600
Dp	541	AAAGCGTCTGACGCTTCCTCCTCGCTGTGAGTCTTATGCTTGGCTGGCGAAAAACAGATCGAT	600
OY	601	CAGCCGTTGCATCTGCGGGGATCACCGAAGCGGATGTGCGGCGACGCGGACAGTAAATCC	660
Dp	601	CAGCCGTTGCATCTGCGGGGATCACCGAAGCGGATGTGCGGCGACGCGGACAGTAAATCC	660
OY	661	GCCATTCGTTAGGTCTGCTGCTGCTGCTGGAAGGCAATCGGCGACACAGCTGCGCGTATCGCTG	720
Dp	661	GCCATTCGTTAGGTCTGCTGCTGCTGCTGGAAGGCAATCGGCGACACAGCTGCGCGTATCGCTG	720
OY	721	GCGCGCCATCCGCTGCGAAGAGATCAAAAGTGGTTTCATATTTTGAATGCGCTGCGATTC	780
Dp	721	GCGCGCCATCCGCTGCGAAGAGATCAAAAGTGGTTTCATATTTTGAATGCGCTGCGATTC	780
OY	781	CGTTGCGGAGGAGTCAACTTCATCGCTGCGCCGACCTGTTGCGGTCAAGAAATTTGATGTT	840
Dp	781	CGTTGCGGAGGAGTCAACTTCATCGCTGCGCCGACCTGTTGCGGTCAAGAAATTTGATGTT	840
OY	841	ATCGCTACGGTTAAACGCGCTGAGAACAGCCTGGAGAGATATCATCTCCGATCGAGCGTT	900
Dp	841	ATCGCTACGGTTAAACGCGCTGAGAACAGCCTGGAGAGATATCATCTCCGATCGAGCGTT	900
OY	901	TTCGATTTCGCGTCCGCGGTGGAATAGGCCCAAGGTGAGGCGCTGGTTTCTACACTCGCGCTC	960
Dp	901	TTCGATTTCGCGTCCGCGGTGGAATAGGCCCAAGGTGAGGCGCTGGTTTCTACACTCGCGCTC	960

OY	1961	ACCGGCGGCACCAAGAAAGCGGCTCTGTATGAAAGATGGCGTGGCAAGACCGTCTGGAC	1020
Db	961	ACCGGCGGCACCAAGAAAGCGGCTCTGTATGAAAGATGGCGTGGCAAGACCGTCTGGAC	1020
OY	1021	AACAACGATATGATGCACCGACTTGGAAAGCAAGCATTTGTCGGAAGCCAGTCAGTCGAC	1080
Db	1021	AACAACGATATGATGCACCGACTTGGAAAGCAAGCATTTGTCGGAAGCCAGTCAGTCGAC	1080
OY	1081	GAAGCGCGTGCATTGACGTTTCAGCAGCTTGAATAATAA	1119
Db	1081	GAAGCGCGTGCATTGACGTTTCAGCAGCTTGAATAATAA	1119
RESULT 4			
ID	AAD31203	AAD31203 standard; DNA: 1119 BP.	
AC	AAD31203:		
DT	31-MAY-2002	(first entry)	
XX	DE	Escherichia coli gcpe gene.	
XX	XX	gcpe gene; methyl-D-erythritol phosphate; MEP; dimethylallyl diphosphate	
KW	KW	transgenic plant; isoprenoid compound; tocopherol; isopentenyl diphosphate	
KW	KW	food; feed source; transfection; single nucleotide polymorphism; SNP;	
KW	KW	oxidative stress tolerance; UV tolerance; transformation; plants; ds.	
OS	XX	Escherichia coli.	
XX	XX	Key	
FT	FT	Location/Qualifiers	
FT	FT	CDS	
FT	FT	1..1119	
FT	FT	/*tag- a	
FT	FT	/product- "Escherichia coli GCPE protein"	
XX	XX	MO200212478-A2.	
PN	XX	14-FEB-2002.	
XX	XX	06-AUG-2001; 2001MO-US24335.	
PR	XX	07-AUG-2000; 2000US-223483P.	
XX	XX	(MONS) MONSANTO TECHNOLOGY LLC.	
PA	PI	Boromati A, Campos N, Rodriguez-concepcion M, Rohmer M, Seeman M;	
PI	PI	Valentin HE, Venkatesh TV, Venkatramesh M;	
XX	XX	WPI: 2002-227151/28.	
DR	DR	P-PSDB; AAE19653.	
XX	XX	gcpe nucleic acid which is an essential gene of the methyl-D-erythritol	
PT	PT	phosphate pathway, encoding a fully defined GCPE protein which is	
PT	PT	useful for increasing levels of tocopherol substrates in plants	
PS	PS	Claim 3; Page 117-119; 155pp; English.	
XX	XX	The invention relates to gcpe nucleic acid molecule, an essential gene	
CC	CC	of methyl-D-erythritol phosphate (MEP) pathway that encodes rice,	
CC	CC	Arabidopsis thaliana or Escherichia coli GCPE protein. gcpe is useful	
CC	CC	for producing a transgenic plant such as Brassica campestris, B. napus,	
CC	CC	canola, castor bean, coconut, cotton, crambe, linseed, maize, mustard,	
CC	CC	oil palm, peanut, rapeseed, rice, safflower, sesame, soybean, sunflower,	
CC	CC	or wheat with an increased isoprenoid (tocopherol) compound level. The	
CC	CC	expression of GCPE protein in organisms increases the level of	
CC	CC	tocopherol substrate such as isopentenyl diphosphate and dimethylallyl	
CC	CC	diphosphate biosynthesis. Transgenic organisms overexpressing GCPE	
CC	CC	protein can nutritionally enhance food and feed sources. Overexpression	
CC	CC	of GCPE protein in transgenic plant may provide tolerance to stresses	
CC	CC	e.g., oxidative stress tolerance such as to oxygen or ozone, UV	
CC	CC	tolerance, etc. gcpe may be used to obtain nucleic acid molecules from	
CC	CC	the same species, and to obtain nucleic acid homologues. gcpe is also	

CC used as or primers. The recombinant vectors are used in plant
 CC transformation or transfection. gcpE an also act as markers capable of
 CC detecting polymorphisms such as single nucleotide polymorphisms (SNPs).
 CC gcpE is also used to determine the level or pattern of expression of
 CC the protein. The present sequence is Escherichia coli gcpE gene.

XX Sequence 1119 BP; 272 A; 279 C; 317 G; 251 T; 0 other:

Query Match 100.0%; Score 1119; DB 24; Length 1119;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCATTAACCAAGGCTTCAATTAACGTAAGAAATCAACACGATATTACGTTGGGAATG 60
 DB 1 ATGCATTAACCAAGGCTTCAATTAACGTAAGAAATCAACACGATATTACGTTGGGAATG 60
 QY 61 CCGATTGGGAGATGCTGCTCCATCGCCCTACAGTCCATGACCAATACGCGTACGACAGAC 120
 DB 61 CCGATTGGGAGATGCTGCTCCATCGCCCTACAGTCCATGACCAATACGCGTACGACAGAC 120
 QY 121 GTGCAAGCAACGGTCAATCAATCAAGGCGTGAACGGCGTATGCTGCTGCT 180
 DB 121 GTGCAAGCAACGGTCAATCAATCAAGGCGTGAACGGCGTATGCTGCTGCTGCT 180
 QY 181 GTATCCGTACCGACGATGAGCGGCGAGAGCGCTTCAAACTCATCAACAGCAGATTAC 240
 DB 181 GTATCCGTACCGACGATGAGCGGCGAGAGCGCTTCAAACTCATCAACAGCAGATTAC 240
 QY 241 GTGCCCGGTGGGTGACATCCACTTCGACTATCGATTGGCCCTGAAATACCGGAATAC 300
 DB 241 GTGCCCGGTGGGTGACATCCACTTCGACTATCGATTGGCCCTGAAATACCGGAATAC 300
 QY 301 GGGCGTCAATGCTGCTGCTTAAACCTGCGCAATATCGTAATGAAGAGCGTATTGCGATG 360
 DB 301 GGGCGTCAATGCTGCTGCTTAAACCTGCGCAATATCGTAATGAAGAGCGTATTGCGATG 360
 QY 361 GTGGTTGACTGTGCGCGCATAAACATTCGATCCGATTGGCGCTTAAACCGCGATCG 420
 DB 361 GTGGTTGACTGTGCGCGCATAAACATTCGATCCGATTGGCGCTTAAACCGCGATCG 420
 QY 421 GTGGAAAGATCTGCAAAAGATATGGCGAAGCGCGCGCGCGCTTGGCGAATCT 480
 DB 421 GTGGAAAGATCTGCAAAAGATATGGCGAAGCGCGCGCGCGCTTGGCGAATCT 480
 QY 481 GGCATCGCTATGTTGATATCATCGATCCGCTGAACCTTGATCAAGTCAAGCAGCTG 540
 DB 481 GGCATCGCTATGTTGATATCATCGATCCGCTGAACCTTGATCAAGTCAAGCAGCTG 540
 QY 541 AAAGCTGTGACGCTCTTCTCGCTGTTGAGTCTTATCGTTTGGTGGCAAAACAGATCAT 600
 DB 541 AAAGCTGTGACGCTCTTCTCGCTGTTGAGTCTTATCGTTTGGTGGCAAAACAGATCAT 600
 QY 601 CAGCCCTTCATCTGGGGATCAACCGAAGCGGCGTGGCGCGCGCGCGCGCATTAATCC 660
 DB 601 CAGCCCTTCATCTGGGGATCAACCGAAGCGGCGTGGCGCGCGCGCGCGCATTAATCC 660
 QY 661 GCCATTGTTTAAAGTCTGCTGCTGTTGAAGCATCGGAGACAGCTGCGCATCGCTG 720
 DB 661 GCCATTGTTTAAAGTCTGCTGCTGTTGAAGCATCGGAGACAGCTGCGCATCGCTG 720
 QY 721 GCGGCGGATCGGTCGAGAGATCAAGTGGTTTGGATATTTTGAATGCTGCTATC 780
 DB 721 GCGGCGGATCGGTCGAGAGATCAAGTGGTTTGGATATTTTGAATGCTGCTATC 780
 QY 781 CGTTGCGAGGGATCAACTTCAATCGCTGCGCACTGTTGGCTGAGGAATTTGATGTT 840
 DB 781 CGTTGCGAGGGATCAACTTCAATCGCTGCGCACTGTTGGCTGAGGAATTTGATGTT 840
 QY 841 ATCGGTACGGTTTAAACGCGTGAAGCAACGCTGGAAGATATCATCTCGATGGAAGCTT 900
 DB 841 ATCGGTACGGTTTAAACGCGTGAAGCAACGCTGGAAGATATCATCTCGATGGAAGCTT 900
 QY 901 TCGATTATCGGCTGCTGCTGTAATGGCCAGGTGAGGCGCTGCTTCTACACTGCGCGTC 960

DB 901 TCGATTATCGGCTGCTGCTGTAATGGCCAGGTGAGGCGCTGCTTCTACACTGCGCGTC 960
 QY 961 ACCGGCGCAACAGAAAAGCGGCTTATGAAGATGGCGTGGCAAAAGACGCTGAGAC 1020
 DB 961 ACCGGCGCAACAGAAAAGCGGCTTATGAAGATGGCGTGGCAAAAGACGCTGAGAC 1020
 QY 1021 AACACGATATGATGACGACGCTGAGACGCAATTCGTGCGAAAGCCAGTCAGCTGAGC 1080
 DB 1021 AACACGATATGATGACGACGCTGAGACGCAATTCGTGCGAAAGCCAGTCAGCTGAGC 1080
 QY 1081 GAAGCGCGTCCGAATTGACGTTCAAGAGTTGAAAATAA 1119
 DB 1081 GAAGCGCGTCCGAATTGACGTTCAAGAGTTGAAAATAA 1119

RESULT 5

AAT42063
 ID AAT42063 standard; DNA; 1830121 BP.

AC AAT42063;
 XX 14-SEP-1999 (first entry)

DE Haemophilus influenzae complete genome sequence.

DE Genome; bacterium; Haemophilus influenzae; computer readable medium;
 KW expression modulating fragment; regulation; gene expression; vector;
 KW organism; open reading frame; ORF; ds.

OS Haemophilus influenzae.

OS WO9633276-A1.

PN 24-OCT-1996.

PD 22-APR-1996; 96WO-US05320.

PF 07-JUN-1995; 95US-0487429.

PR 21-APR-1995; 95US-0426787.

PR 07-JUN-1995; 95US-0476102.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (UYJO) UNIV JOHNS HOPKINS.

XX Adams MD, Fleischmann RD, Smith HO, Venter JC, White O;

XX WPI; 1996-485782/48.

XX Haemophilus influenzae Rd genome recorded on computer readable

PT medium - useful for identifying commercially important nucleic acid

PT fragments by homology searching

XX Claim 1: Page 77.2-77.1091; 1291pp; English.

PS This sequence represents the complete genome sequence of the bacterium
 CC Haemophilus influenzae strain Rd. The invention relates to a computer
 CC readable medium (CRM) having recorded upon it the complete H. influenzae
 CC nucleotide sequence (I), a representative fragment of (I) or a nucleotide
 CC sequence at least 99% identical to (I). By providing the full-length
 CC genomic sequence in a computer readable form, it is possible to identify
 CC commercially important nucleic acid fragments and expression modulating
 CC fragments (EMFs) of the Haemophilus genome. The EMFs can be used to
 CC regulate the expression of a nucleic acid molecule. Vectors and altered
 CC organisms comprising the predicted ORFs can be used to produce any of the
 CC polypeptide fragments of the H. influenzae Rd genome.

XX Sequence 1830121 BP; 567399 A; 350615 C; 347389 G; 564036 T; 682 other;

Query Match 55.6%; Score 622; DB 17; Length 1830121;
 Best Local Similarity 74.8%; Pred. No. 3.4e-169;
 Matches 794; Conservative 0; Mismatches 265; Indels 3; Gaps 1;

QY	10	CAGCGTCGAATTCACACGTAAGAAAATCACACGCGTATTTACGTTGGGAATGCGGCATGGC	69
Db	390972	CAGCCACATTTAAACGCGTCGTGAATCGACAAAATTTATGTGGGAAATGTACCAATTTGCT	391031
QY	70	GATGGTGCTCCCATCGCCGTCACAGTCATGCATAGCAATACGCGTACGACAGACGTCGAAGCA	129
Db	391032	GGGGATGGGCGCCATTGTGCCGTGCATCAATGACAAATATCTCCACACTGATGTGGAAAGCG	391091
QY	130	ACGGCATTCAAATCAACGCGCGTGGAAAGCGCTGGCGCGTGATACGTCGCTGTATCCGCTA	189
Db	391092	ACAGTTGCTCAAAATTTAAATTCATTGAACGTTGTGGTGCAGATATATGTCGTGTATCGTT	391151
QY	190	CCGAGCATGGACGCGCGCAGAAAGCGTTTCAACATCAATCAACAGCAGGTAAAGTTCGCGCTG	249
Db	391152	CCAAACATGGATGCGTGGGGAAGCAATTTAAACAATTTAAACACACAGTGAATGTTCCGCTC	391211
QY	250	GTGGCTGACATTCACCTTGACTATGCGATTTGCGCTGAAAGTACGCGAATACGCGCTCGAT	309
Db	391212	GTAGCAGATTTATTCATTTCTGCAATTCGTATCCGCTTTAAAGTCGACAAATATGAGTGGAT	391271
QY	310	TGTCGCGATTTAAACCCGCGCAATATGGGTAAATGAAGCGGTATTCGATAGTGGTTGAC	369
Db	391272	TGTTACGTTATCAATCTCGGCAACATTTGGTGTGTAAGATGCGCGTCGTCCGCTGTTGGAT	391331
QY	370	TGTGCGCGCGATTAATAAACATTCCTCGATCCGTAATTTGGCGCTTTAAGCCGCGATCGCTGGA	429
Db	391332	TGTGCGCGAGCAAAAATATTTCCGATTCGTATGTGTGTAAATGACGCGCTCTTAGAATAAA	391391
QY	430	GATCTGCAAGAAAGATATGGCGAAACCGACCGCGCAGCGCTTCTGTGAATCTGCCATGCGT	489
Db	391392	GATTTTGCAAGAAAAATATGGGGAACCAACGCGCAACCCCTGTTAGAAATCCGATTTGCGT	391451
QY	490	CATGTTGATTCATCTCGATCGCGCTCAACTCGATCGATAGTTTCAAAGCACGCGGAAGCGTCT	549
Db	391452	CATGTAGAAATTTCTAGATCGGCTTAACTTGATGATAGTTTAAAGTGAAGCGTAAAGCGCTCC	391511
QY	550	GACGTCCTTCCTCGCTGTTGAGTCTTATGCTTTATGCTGTGCGCAAAACAGATCGATCAGCCGTTG	609
Db	391512	GATGATATCTTGAAGCGGTTGAATCTTACGTTTACTGCGCTAAAGCAATTTAAACAGCCTTTA	391571
QY	610	CATCTGGGGATCACGGAAGCGCGGTGGGCGCGACGCGGGGCACTAAATCCGCGCATTTGCT	669
Db	391572	CATTTAGGCATTTACAGAAAGCAGGTGGCGCACGCGCTGTGCAGTAAATGTCAAGTGGGT	391631
QY	670	TTAGGTCGTGCTGTCTGTAAGGAGCATCGGCGACACGCTGGCGCTATTCGCTGGCGGCGCAT	729
Db	391632	TTAGGAATGTTATTAGCTGAGAGGCGATTGGCGATACACTACGCGCTGTTGGGCGCAAT	391691
QY	730	CCGGTCGAAGATCAAAAGTCGGTTTGATATTTTGAATTCGCTGGCGTATCCGTTGGCGA	789
Db	391692	CTGTAGAGGAAATCAAAAGTCGGTTTGATATTTTGAATCTTACGATTCCTTCACAA	391751
QY	790	GGGATCAACTTCATCGCGCTGCCGCACTGTCCGCTGAGAGATTTATGTTATGATTCGCTAGG	849
Db	391752	GGATTTAACTTTATTTGCTGGCCCACTGTCTTCGCAAAATTTGATGTAAATCGGTACAA	391811
QY	850	GTTAACCGCTGGAGCAACGCGCTGGAAGATATCATCTCCGATGGAAGCTTTGCATTATC	909
Db	391812	GTAATGCGCTAGAAACAACGCGCTTGAAGATATTTATACACCAATGATGATATTTATTC	391871
QY	910	GGCTGCGGTGGATGGCCGACGCTGAGGCGCTGGTTTCTACATCGGCGTCAACGCGCGCGC	969
Db	391872	GCTTGTGTAGTATGGTCTCTGGCGAAGCACTGCTCTTCGATCTGGCGGTAAACGGCGGCT	391931
QY	970	AACAAGAAAAAGCGCTCTATGGAAGATGGCGTGGC---AAAGACCGTCTGCAACAAC	1026
Db	391932	AACAAAAAAGCGGTTATTTATCTTTAGCAGGAAGCGCAAAAAAGAGCGTTTGTATACGAA	391991
QY	1027	GATATGATTCGACCGCTGCGAAGACGACGATTCGTGCGCAAGCC	1068
Db	391992	GATATGCTGAACCAATTAGGAAGAAAAATTCGTGCGCAAACTC	392033

[illegible]

QY 755 TCATATTTTGAATTCGCTGCGTATCCGTTCCGAGGATCATCTGCTCCCTGCCGA 814
DB 2615 TCATATTTTGAATTCGCTGCGTATCCGTTCCGAGGATCATCTGCTCCCTGCCGA 2674
QY 815 CCTGTTCCGCTGAGGATTTATGTTATCGGTAGCGTTAACGCGCTGAGACAGCGCTGG 874
DB 2675 CCTGTTCCGCTGAGGATTTATGTTATCGGTAGCGTTAACGCGCTGAGACAGCGCTGG 2734
QY 875 AAGATATCATCTCCGATGAGACGTTTGTATTCGCTGCTGCTGATGATGCGCCAGGTG 934
DB 2735 AAGATATCATCTCCGATGAGACGTTTGTATTCGCTGCTGCTGATGATGCGCCAGGTG 2794
QY 935 AGGCGTGGTTTCTACACTCGCGCTGACCGCGGACACAGAAAAGCGCTCTATGAG 994
DB 2795 AGGCGTGGTTTCTACACTCGCGCTGACCGCGGACACAGAAAAGCGCTCTATGAG 2854
QY 995 ATGGCGTCCGCAAGACCGCTGTGACACACACATATGATGACCGTGGAGACGCA 1054
DB 2855 ATGGCGTCCGCAAGACCGCTGTGACACACACATATGATGACCGTGGAGACGCA 2914
QY 1055 TTCGTGCGAAGCCAGTCAGCTG 1077
DB 2915 TTCGTGCGAAGCCAGTCAGCTG 2937

RESULT 7
ABA92787

ID ABA92787 standard; DNA: 640681 BP.

AC ABA92787;

DT 27-MAR-2002 (first entry)

DE Buchnera sp. genomic DNA SEQ ID NO:1.

KW Buchnera; cockroach-symbiotic bacterium; cockroach extermination;

KM circular; ds.

OS Buchnera sp.

PN JP2001292771-A.

PD 23-OCT-2001.

PF 07-APR-2000; 2000JP-0107160.

PR 07-APR-2000; 2000JP-0107160.

PA (RIKA) RIKAGAKU KENKYUSHO.

DR WPI: 2002-126043/17.

XX A genomic DNA of cockroach-symbiotic bacterium

PS Claim 1: Page 16-230; 237pp; Japanese.

XX The present invention describes a gene (I) derived from Buchnera sp.
XX containing the DNA (a) or (b), (a) has a fully defined base pair
XX sequence selected from a table of sequences found in the Buchnera sp.
XX genomic DNA of ABA92787 given in the specification or is a DNA selected
XX from complementary DNA sequences, and (b) is a DNA which hybridises with
XX the DNA (a) and encodes a protein. Also described are: (1) a recombinant
XX vector (II) containing (I); (2) a transformant (III) containing (II);
XX (3) a genomic DNA of Buchnera sp. containing the sequence given in
XX ABA92787; (4) a plasmid derived from Buchnera sp. containing DNA (c) or
XX (d), (c) is a DNA containing a fully defined sequence given in ABA92788
XX or ABA92789 and (d) is a plasmid which hybridises with a DNA; and (5) a
XX method for the preparation of a protein in which (III) is cultured and
XX the expression protein of the objective protein is collected from the
XX resultant culture. The DNA is useful for developing agricultural
XX chemicals for exterminating cockroaches. The present sequence represents
XX the specifically claimed Buchnera sp. genomic DNA sequence, from the

CC present invention.

SO Sequence 640681 BP; 237522 A; 83822 C; 84757 G; 234580 T; 0 other;

Query Match 36.1%; Score 403.6; DB 24; Length 640681;

Best Local Similarity 62.2%; Pred. No. 7.2e-106;

Matches 653; Conservative 0; Mismatches 394; Indels 3; Gaps 1;

QY 19 ATTCAACGTAAATCAACACGATATTACGTTGGGANTGCGCGATGCGTGCCT 78
DB 314293 ATCATAGAAAGAAATTCGATCGTATTTATGTGAAAAGTCGATTTGGCAATTAATCG 314352
QY 79 CCGATGCGGCTACAGTCCATGACCAATAGCGTACGACAGCGTGAAGCAACGCTCAT 138
DB 314353 CCAATATCAGTTCAATCTATGACAAATATCTGCTACTATTAATCTGGAAGCTATTAAT 314412
QY 139 CAATTCAGGCGCTGGAACGGGTTGGCCTATATCGTCCGTGTTCCGTACCCGACGAT 198
DB 314413 CAATCTTAGAGTTACAAAAGATAGAGTATGTTGCTGATTTCTATACCCAAATTTA 314472
QY 199 GACGCGGAGAGCGGTTCAAACTCATCAACAGCAGTTAACGTCGCGCTGCTGCTGAC 258
DB 314473 AAAGCTGAGAAATCATTTCAAAAAGAAATATAAAGCAACAAATGTTCCATGATTCAGAT 314532
QY 259 ATCCACTTGAATATGCAATTCGCTGCTGAAAGTACGCGAATAGCGGCTGATTTGCTGCT 318
DB 314533 ATACATTTTGAATACAGATTTAGCTTTACAGCTATTAATATGTTGTCAGATTTGTTAAGA 314592
QY 319 ATTAAACCTGGCAATATGCGTAAATGAAGCGTATTCGCAATGCTGCTGACCTGTCGCGG 378
DB 314593 ATTAACTCTGGAAATATTGGAATTAAGAAAGATATCAGAAATCATTTCTTACCAANA 314652
QY 379 GATAAAACATTCGATCCGATTCGCTTAAAGCGCGGATGCTGGAAGAAATGTCGA 438
DB 314653 GATGAAATATTTCCATTTTCGATTTGCTTATGCTGATCTTTTGAAGAAATATTTA 314712
QY 439 GAAAAGTATGGCGAAGCAGCAGCGCGCTGCTGTAATCTGCCATCGCTCATGTTGAT 498
DB 314713 AAAAATATTAAATATCTACCTCAAGTACAGTATGATGATGATGATGATGATGAT 314772
QY 499 CATCTCGATCGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 558
DB 314773 TACTTGTATGCTTTAAATTTTAAATTTTAAAGTTAGTAAAGCTGATGATTTATTT 314832
QY 559 CTCGCTGTTAGTCTTATGCTTTGCTGCAAAACAGATCATCACCGCTGATCTGATCTG 618
DB 314833 TTAGCTATTTGATCATATTCGATGATTTAGGAAAGAAATTTACACACCTTTGATTTG 314892
QY 619 ATCACCAGAACCGGCTGCTGCGCGACAGCGGCGAGTAAATCCGCAATTTGTTAGCTG 678
DB 314893 ATAACTGAATCCGCTGCTGTTAAGGAATGAAACAGTTAAATCATCTATGATTTCTTTA 314952
QY 679 CTGCTGCTGAAGGATGAGGCGGACAGCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 738
DB 314953 TTATTATTGAAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 315012
QY 739 GAGATCAAAAGCTGTTGATTTGATTTGAAATGCTGCTGATCTGCTGCGAGGATCAAC 798
DB 315013 GAAATGAAAGTATGATTTGATTTGAAATGCTGCTGATCTGCTGCGAGGATTTAT 315072
QY 799 TTGATGCGCTGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 858
DB 315073 TTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 315132
QY 859 CTGAGGCAAGCGCTGAGGATGATCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 918
DB 315133 CTAGGAAAGAAATCTAGGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 315192
QY 919 GTGATGCGCGGAGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 978
DB 315193 GTTATGGAATAGGATGATCTGATTAATAGCACTTTAGGCTGCTGCTGCTGCTGCTGCT 315252
QY 979 AGCGGCTCTATGAGATGGCT---GGCAAGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1035

Db 315253 AGTCCATTATATGAGACGGGTAGACAAAAAGAAAAATTAATAAGAAATTATA 315312
 Oy 1036 GACCAGCTGGAGCAGCAGATTCTGCGGAAA 1065
 Db 315313 GAAAAAATGGAAATTAATAATTGAAAAA 315342

RESULT 8

AAF28536 ID AAF28536 standard; DNA: 33140 BP.

AAF28536; AC

04-APR-2001 (first entry) DE

Genomic fragment #23. DE

Genomic library; bacteriophage; human upper airway; otitis media; sinusitis; KM

bronchopulmonary; endocarditis; meningitis; ss. KM

Moraxella catarrhalis. OS

WO200078968-A2. PN

28-DEC-2000. PD

16-JUN-2000; 2000MO-US16649. PF

18-JUN-1999; 99US-0140121. PR

(INCY-) INCYTE GENOMICS INC. PA

Lagace RE, Patterson C, Berg KL; PI

WPI; 2001-041427/05. DR

Genomic library for identifying diagnostic and therapeutic PT

compositions, and for identifying virulence factors, regulatory PT

elements and drug targets, comprises Moraxella catarrhalis nucleic PT

acids - PS

Claim 1; Page 191-199; 545pp; English. PS

The present invention relates to a Moraxella catarrhalis genomic library CC

comprising of a combination of 41 nucleic acid molecules (see CC

AAF28514-AAF28554). The library has a number of uses described in the CC

specification e.g. is useful for identifying diagnostic and therapeutic CC

compositions. M. catarrhalis (Branhamella catarrhalis) is a large CC

aerobic, gram-negative diplococcus, normally found among the bacterial CC

flora of human upper airways. M. catarrhalis is known to cause acute, CC

localised infections such as otitis media, sinusitis and bronchopulmonary CC

infection and life-threatening, systemic diseases including endocarditis CC

and meningitis. XX

Sequence 33140 BP; 9472 A; 6496 C; 7384 G; 9787 T; 1 other; SQ

Query Match 34.2%; Score 382.6; DB 22; Length 33140; Best Local Similarity 59.2%; Pred. No. 2.6e-100;

Matches 652; Conservative 0; Mismatches 449; Indels 0; Gaps 0;

1 ATGCATTAACCAAGGCTCCATTTCAAGTAGAAAAATCAACAGTATTTCAGTTGGGAATGTG 60

21735 ATGTCAATACACAATCGATTAAAGCGCCCTAACCAAAAAATCTATGTGGCAATGTC 21794

61 CCGATTGGCGATGCTGCCATCGCCGTAACATCCATGACCAATACGCGTACGACAGAC 120

21795 GCCATGGCGGCGCATCACCACATCAGCTGCAAGATATGCAAAATCAACACCTGTGAT 21854

121 GTGCAAGCAACGCTCAATCAATCAAGCGCGTGAACGCGTGGCGGTATGCTGCGGT 180

21855 ATTGATGCAACAGTGTGACAAATTCAGCGATGCGTGAACGCGGTGGCGGTGATGCGGT 21914

Oy 181 GTATCCGTACCGACGATGACGCGGACAGACCGTTCAAACTCAACACAGACGTTAAC 240

21915 GTATGACACCAACGATGAGAGAGCGTGTGCTTCTGATGATTAATAAAAGCGCTATCC 21974

Oy 241 GTGCGGCTGGTGGTGCATATCCACTTGCATATCGCATTTGGCGTGAAGACGGAATAC 300

21975 ATTCCACTCAATGGCGATATTCATTTGATTCACAAAATCGCATTTGACGCGGATGTG 22034

Oy 301 GCGCTGATGTGTGCGATTAACCGGCAATATGCGTAAAGAGCGGATATTCGATG 360

22035 GTGCAAGACTGCTGCGATATCAATCAGTATATTTGCAATGATCAAAAAGTCAAAAGAA 22094

Oy 361 GTGGTGTACTGTGCGCGCATTAATAACATTCGATTCGATTTGCGTAAACGCGGATCG 420

22095 GTGGTGTACAGGCGACCGCATTAATGTCGCGATTTGCGTATGCGTATGCGATGCA 22154

Oy 421 CTGGAATAAGATCTGCAAGAAAGTATGCGCAACCGACGCGCGCTTGTGGAATCT 480

22155 CTGGAATAAGATTTACAAAAAATATCGAAGACCGACGCGGATGCGATGGAATCG 22214

Oy 481 GCCATGGCTATGTTGATCATCTGATCGCGTGAATTCGATCATGATCAAGTCAGCGTG 540

22215 GCGTTACGCCACATTTATTTAGAAAAATCTTATTTAGAAATTTAAATCTCAGTC 22274

Oy 541 AAAGCGTGTACGCTCTCTCGCTGTGAGTCTTATGCTTGTGCGCAAAACAGATGAT 600

22275 AAAGCGTGTACGCTCTCTCGCTGTGAGTCTTATGCTTGTGCGCAAAACAGATGAT 22334

Oy 601 CAGCGTGTGATCTGGGATCACCAGAACCGGTGTGTGCGGACGCGGACGATCAATCC 660

22335 AATCCATTTGATCTGGGATCACCAGAACCGGTGTGTGCGGACGCGGACGATCAATCC 22394

Oy 661 GCCATTTGTTAGTGTGCTGTGTGAGGACATCGCGACACGCGGATGCGGATGCGTG 720

22395 GCCATTTGTTAGTGTGCTGTGTGAGGACATCGCGACACGCGGATGCGGATGCGTG 22454

Oy 721 GCGGCGCATCGGTGAGAGATCAAGTGTGTTGATTTGAAATCGTGTGATC 780

22455 GCGGCGCATCGGTGAGAGATCAAGTGTGTTGATTTGAAATCGTGTGATC 22514

Oy 781 CTTTCGCGAGGATCAACTTCATCTGCGTGTGCGGACCGCTGTGCGTCAAGATTTGATGT 840

22515 CTTTCGCGAGGATCAACTTCATCTGCGTGTGCGGACCGCTGTGCGTCAAGATTTGATGT 22574

Oy 841 ATCGGTACGCTTAAACGCGTGTGAGCAACGCGTGTGAGATTCATCTCGATGACGTT 900

22575 ATCGGTACGCTTAAACGCGTGTGAGCAACGCGTGTGAGATTCATCTCGATGACGTT 22634

Oy 901 TCGATTATCGGCTGCGGTGTGAAATGCGGCGGTGAGGCGGTGTTCTACACTGCGGCTG 960

22635 TCGATTATCGGCTGCGGTGTGAAATGCGGCGGTGAGGCGGTGTTCTACACTGCGGCTG 22694

Oy 961 ACCGCGGCGCAACAAGAAACGCGCTCTATGAAGATGTGCGTGTGCGCAAAAGCCGTGTGAC 1020

22695 ACCGCGGCGCAACAAGAAACGCGCTCTATGAAGATGTGCGTGTGCGCAAAAGCCGTGTGAC 1026

Oy 1021 AACACGATATGATTCGACCGCTGTGAAAGCAACGCTGTGCGCAAGCACTGACGTGAC 1080

22755 AACACGATATGATTCGACCGCTGTGAAAGCAACGCTGTGCGCAAGCACTGACGTGAC 1086

Oy 1081 GAAAGCGGTGGAATTCAGCTT 1101

22815 GAAAGCGGTGGAATTCAGCTT 22835

RESULT 9

AAS90066 ID AAS90066 standard; cDNA: 1857 BP.

AAS90066; AC

13-FEB-2002 (first entry) DT

XX

DE DNA encoding novel human diagnostic protein #25870.
 XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX Homo sapiens.
 OS WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YT;
 PI WPI: 2001-639362/73.
 DR P-PSDB; ABG25879.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX Claim 1: SEQ ID NO 25870; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful for treating
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX Sequence 1857 BP; 463 A; 472 C; 530 G; 392 T; 0 other:
 Query Match 31.8%; Score 356; DB 23; Length 1857;
 Best Local Similarity 97.9%; Pred. No. 4.2e-93;
 Matches 466; Conservative 0; Mismatches 0; Indels 10; Gaps 10;
 QY 644 GCGGGGCAATTAATCCGCGATGTTAGTGTCTGCTGTGAAGCATGGGGACA 703
 DB 512 GCGGGGCAATTAATCCGCGATGTTAGTGTCTGCTGTGAAGCATGGGGACA 570
 QY 704 CGGTGGCGATGCGTGGGGCGATCCGCGTGAAGATCAAGTCGGTTGATATTT 763
 DB 571 CGGTGGCGATGCGTGGGGCGATCCGCGTGAAGATCAAGTCGGTTGATATTT 629
 QY 764 TGAATCGCTGATCCGTCGAGGATCAATTCATGCGCTGCCGACCTGTTGCG 823
 DB 630 TGAATCGCTGATCCGTCGAGGATCAATTCATGCGCTGCCGACCTGTTGCG 688
 QY 824 GTGAGCAATTTGATGTTATCGGTAACGCGTGGAGCAACGCTGGAAGATATCA 883
 DB 689 G-CAGGAATTTGATGTTATCGGTAACGCGTGGAGCAACGCTGCG-AGATATCA 746

QY 884 TCACCTCCGATGAGACGTTTGCATTTATCGGCTGGTGAATGCCAGTGAGGCGCTGG 943
 DB 747 TCACCTCCGATGAGACGTTTGCATTTATCGGCTGGTGAAT GCCAGGTGAGGCGCTGG 805
 QY 944 TTTCTACACTGGCGCTTCACCGCGGCAACAGAAAAGCGCTCTATGAAGTGGCGTGC 1003
 DB 806 TTTCTACACTGGCGCTTCACCGCGGCAACAA-AAAAGCGCTCTATGAAGTGG-GTGC 863
 QY 1004 GCAAGACCGCTGAGCAACAGATATGATGACACGAGTGAAGACGATTCGTCGA 1063
 DB 864 GCAAGACCGCTGAGCAACAA-GATATGATGACACGAGTGAAGACGATTCGTCGA 922
 QY 1064 AAGCAGTACGCTGAGCAAGCGCGTCAATTTAGCTTCAGAGGTTGAAAATTA 1119
 DB 923 AAGCAGTACG-AGCAGCAAGCGCGTCAATTTAGCTTCAGAGGTTGAAAATTA 977
 RESULT 10
 ABK74401
 ID ABK74401 standard; DNA; 1083 BP.
 XX AC ABK74401;
 XX 13-AUG-2002 (first entry)
 DE Bacillus licheniformis genomic sequence tag (GST) #1692.
 XX Differential gene expression; genomic sequenced tag; GST;
 KM altered culture condition; environmental stress;
 KM physiological provocation; ds.
 OS Bacillus licheniformis.
 XX WO200229113-A2.
 XX 11-APR-2002.
 PD 05-OCT-2001; 2001WO-US31437.
 PF 06-OCT-2000; 2000US-0680598.
 PR 27-MAR-2001; 2001US-279526P.
 XX (NOVO) NOVOZYMES BIOTECH INC.
 PA (NOVO) NOVOZYMES AS.
 XX Berka R, Clausen IG;
 PI WPI: 2002-416684/44.
 DR Monitoring differential expression of several genes in first Bacillus
 PT cell relative to expression of same genes in one or more second
 PT Bacillus cells, by using substrate containing Bacillus genomic,
 PT sequenced tag array -
 XX Claim 4: SEQ ID NO 1692; 200pp; English.
 PS The invention describes a method of monitoring differential expression of
 CC genes in a first Bacillus cell relative to expression of the genes in
 CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
 CC isolated from Bacillus cells to a substrate containing array of Bacillus
 CC genomic sequenced tags (GST), examining the array, and determining
 CC relative gene expression by an observed hybridisation reporter signal of
 CC a spot in the array. The method is useful for measuring the expression of
 CC genes in a first Bacillus cell relative to expression of the same genes
 CC in one or more second Bacillus cells. The method is useful for monitoring
 CC global expression of several genes from a Bacillus cell, discovering new
 CC genes, identifying possible functions of unknown open reading frames and
 CC monitoring gene copy number variation and stability. Monitoring changes
 CC in expression of genes may be used to provide a representation of the way
 CC in which Bacillus cells adapt to changes in culture conditions,
 CC environmental stress or other physiological provocation. Extensive
 CC follow-up characterisation is unnecessary, when one spot on an array

CC equals one gene or one open reading frame, since sequence information is available. This sequence represents a genomic sequence tag (GST) used in the method of the invention.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 1083 BP; 314 A; 278 C; 265 G; 226 T; 0 other;

Query Match 26.2%; Score 293.4; DB 24; Length 1083;

Best Local Similarity 56.6%; Pred. No. 5.3e-75;

Matches 343; Conservative 0; Mismatches 416; Indels 0; Gaps 0;

```

QY 22 CAACGTAGAAATCAACAGGATTTAGCTTGGGAATGCGGATGGGATGGTGTCTCC 81
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 21 CATGTACAAAAGCGGCGGCTTAAGTGGACCTTTACAAATAGCGGCAATACGAA 80
QY 82 ATGCGCGTACAGTCCATGAGCAATACGCGTACGACAGAGTGAAGCAACGTCATCAA 141
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 81 GTCTCATTTCAACAGCATAGAACAAAACACATGACGTTGAAGCAACCGTCGCCGAA 140
QY 142 ATCAAGCGCGTGAAGCGGCGGCGATGATGTCCTGATACCGCATGAGAC 201
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 141 ATCAAGAGCTCGCGGAGAGGATGTCATTAATCGTCGCTGCTGTGATGAACGG 200
QY 202 GCGGAGAGCGCTCAAAAGCATCAACAGAGTAACTGCGCGTGTGGTGCATC 261
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 201 GCTCCGAGCGCATTCAGAGATCAAAAAGGAGATTCATCCCTCTGTGTGATATT 260
QY 262 CACTTCAGTATGCGATTTGGCTGAAGTAGCGAATACGCGCTGATTTGTCTGATTT 321
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 261 CATTTCACATTAATTTGGGATTAAAGCATCAAGCGGAGCGGATTAATCCGATC 320
QY 322 AACCTTGCAATATGCTTAAAGAGCGATTCGATGGGCTGTGATGCGCGGAT 381
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 321 AATCCGGGTAACTCGCGCGCGGAGGAAAGTTTAAACGCTGTCATCAACGCAAGAA 380
QY 382 AAAACATTCGATCGGTATGGCGTTAACCGCGATCGTGAAGAAAAGATCTGCAAGAA 441
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 381 AAGGCGATTCGATCGGATCGGCGTCAATGACAGCTCTCTGAAAAAAGCATCTTGAG 440
QY 442 AAGATGGCAACGAGCGGAGCGGCTGTGTAATCTGCCATGCTCATTTGATCAT 501
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 441 AAGATGGGTAACTCGCGCGCGGAGGAGCGGATGTCGAAAGCCCGCCACCATTAAT 500
QY 502 CTGATGCGCTGAATCTGATCAAGTCAAGTCAAGTCAAGGAGGAGGAGGAGGAGGAG 561
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 501 CTGAGAGATCTGATTCGAGATATCATGTCAGCATGAGAGGCGTGTGATGTAACCTG 560
QY 562 GCTGTGAGTCTTATGCTTGTGCAAAACAGATCGATCAGCGGTTGATCTGGGATC 621
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 561 GCGATTTAGGCAATGAAAGAGGCGGCTTAAGCTTTCGATTCGCTTCAATTTAGAGCAT 620
QY 622 ACCGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 681
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 621 ACCGATTCGGAAGCGCTGTGCGGAGTACAGTGAAGCCCGCGGCTGTGCGGCGCATC 680
QY 682 CTGTCTGAAGCATCGGCGGAGCGTGCATGCTGCGGCGGAGTCCGCTGAGAA 741
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 681 CTTTCAAAAGAAATCGGCAATACATTTGCGGATTTCTTAAGCGGAGCCCGGTCGAGAA 740
QY 742 ATCAAGTCTGTTTCAATTTTGAATCGCTCGATTCGCTTGGCGAGGATCAATTC 801
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 741 GTAAAAAGTCCAGAGGAGCTGAATCTTTGCGGCTCTCCATTCGCGCAACATTTG 800
QY 802 ATGCGCTGCGCGAGCTGTGGCGTCAAGATTTGATGTTTTCGCTAGCGGATCAATTC 861
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 801 ATTCTCGCCGAGCTTGGCGGCGATGCAATGATTTGATTCGATTCGCAATGAAATTC 860
QY 862 GAGCAAGCGCTGGAAGATATGATCCTGATGAGGAGGATTTTCGCTGCGGCTGAGT 921
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 861 GAAATTTACATCGCAAAAGAGGCGCGATCAAGATTTGCGGCTCTGCGGCTGAGT 920

```

```

QY 922 AATGCCAGGTAGAGCGCTGTTTCTACACTCGGCGTACCGCGGCAACAGAAAG 980
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 921 AACGCTCCGAGAGAGCGCGGAGCGGATTCGAGATCGCGGCGCACGCGGTAATG 979

```

RESULT 11

AB070939 standard; DNA: 6157 BP.

AB070939;

29-AUG-2002 (first entry)

Listeria monocytogenes 4b contig DNA sequence #881.

Antibacterial; Listeria; food contamination; mutational analysis; infection; ds.

Listeria monocytogenes 4b.

W0200228891-A2.

11-APR-2002.

04-OCT-2001; 2001WO-FR03061.

04-OCT-2000; 2000FR-0012697.

(INSP) INST PASTEUR.

(CNRS) CNRS CENT NAT RECH SCI.

Kunst F, Glaser P;

WPI; 2002-332479/37.

New genomic sequences from Listeria species, useful for detection, treatment and prevention of infection, also related polypeptides, antibodies and modulators

Claim 14; SEQ ID 3752; 180pp; French.

The present invention relates to nucleic acid sequences (AB067188-AB071212) from Listeria sp. The sequences are useful as probes and primers for identification and/or detection of Listeria (e.g. as contaminants in foods, or mutational analysis) and for analysis of gene expression. Proteins encoded by the nucleic acid sequences can be used to screen for compounds that modulate gene expression, replication and pathogenicity of Listeria (potential therapeutic agents), also for treating infections by Listeria, and are useful as immunogens in anti-Listeria vaccines.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

Sequence 6157 BP; 1731 A; 1065 C; 1209 G; 2150 T; 2 other;

Query Match 22.8%; Score 254.6; DB 24; Length 6157;

Best Local Similarity 53.7%; Pred. No. 2.2e-63;

Matches 527; Conservative 0; Mismatches 454; Indels 0; Gaps 0;

```

QY 18 AATTCAACGTAGAAATCAACACGATTTAGCTTGGGAATGCGGATTTGGCGATGTC 77
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 3116 AATCTTCCGGAAGAACACTCGGCGCATCCAAAGTAGGTAATTTAATTTGTTGTA 3175
QY 78 TCCCATCGCGGTACAGTCCATGCAATAGCGGTACGACAGAGCTGAAGCAACGTTCA 137
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 3176 GGAATTAATCTATCAAGAGTACTACTACAAAGACATGATGTGAAGCAACAGTAC 3235
QY 138 TCAATCAAGGCGCTGGAACGCGTGTGCGCTGATTCGCGGATTCGCTTACGACGAT 197
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 3236 GGAATTAACCGATTGAAAGAGGTGGTGTGATGATTTGCGAGTGTCTTCTGATGA 3295

```


Db 492 AATTCGAGAGATTACAGATTATGATATCATTTGTTCTTGAAGGCTCTGATGAA 551
 QY 558 CCTGCTGTTGAGTCTTATGTTGCTGCGCAAAACAGATCGATCGCGTTCATCTGG 617
 Db 552 TTATGACATGACCTTATGATTAAGTACGCGCATTTATATCTCTGCACTCGG 611
 QY 618 GATACACGAGCCGCTGGTGGCGGACGGGCGAGTAATAATCCCATTTGTTAGTCT 677
 Db 612 AATTACAGAACTGCTACAGAAATTTGCTGGAGAAATTAAGAATGCTGCTGTTAGAGC 671
 QY 678 GCTCTGCTGCAAGGAGGAGGCGGACGCTGCGGTATTCGCTGGCGGCGCATCGGTCGA 737
 Db 672 GATACCTAGTTTGGGCACTTGGAAATACATTAACGATATCTTTGAGTCTGATCTGGA 731
 QY 738 AGAGATCAAGAGTGGTTGCTGATTTTGAATCGCTGATCCGTCGCGAGGATCA 797
 Db 732 AGAAATTAAGTGGCCGCGGAGTTTAAATATTTGTCCTTCCGATGCCGCAAT 791
 QY 798 CTTCATGCGCTGCGCGACGCTGCTGCGTCAAGAAATTTGATGTTATCGGTAAACGC 857
 Db 792 GCTTATCTCTGCTGCTGCTGCGGTCGAAATAGAGATTTAATGCTATCGCTAATGA 851
 QY 858 GCTGAGCAAGCGCTGGAATATCATCTCCGATGAGGCTTTCGATTAATCGGCTCGT 917
 Db 852 AGTGGAAATTTACATACCAATGATGAAGTTCGATTAAGTACCGCTGCTGCTGCG 911
 QY 918 GGTGATGCGCCAGTGAAGGCTGTTTCTACACTGCGGTCACCGCGGCAACAA 977
 Db 912 GGTCAACGCGCTGGAAGACGCTCGCAAGCGATATGCAATTTGCTGTTCAACGAGGA 971
 QY 978 AAGGCGCTCTATGAGATG 998
 Db 972 AGGCTTCTTTTACATG 992

RESULT 13 AB069993

AB069993 standard; DNA; 3107 BP.

AC AB069993;

DT 29-AUG-2002 (first entry)

DE Listeria monocytogenes EDE DNA sequence #205.

KW Antibacterial; Listeria; food contamination; mutational analysis;

OS Listeria monocytogenes EDE.

PN MO200228891-A2.

PD 11-APR-2002.

PF 04-OCT-2001; 2001MO-FR03061.

PR 04-OCT-2000; 2000FR-0012697.

PA (INSP) INST PASTEUR.

PA (CNRS) CNRS CENT NAT RECH SCI.

PI Kunst F, Glaser P.

DR WPI; 2002-332479/37.

PT New genomic sequences from Listeria species, useful for detection,

PS treatment and prevention of infection, also related polypeptides,

CC Claim 16; SEQ ID 2806; 180pp; French.

CC The present invention relates to nucleic acid sequences

(AB067188-AB071212) from Listeria sp. The sequences are useful as probes

CC and primers for identification and/or detection of Listeria (e.g. as
 CC contaminants in foods, or mutational analysis) and for analysis of
 CC gene expression. Proteins encoded by the nucleic acid sequences can be
 CC used to screen for compounds that modulate gene expression, replication
 CC and pathogenicity of Listeria (potential therapeutic agents), also for
 CC treating infections by Listeria, and are useful as immunogens in
 CC anti-Listeria vaccines.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pcr-sequences.
 CC XX
 SQ Sequence 1107 BP; 348 A; 178 C; 255 G; 326 T; 0 other;

Query Match 22.5%; Score 251.4; DB 24; Length 1107;

Best Local Similarity 53.5%; Pred. No. 8.8e-63;

Matches 525; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

QY 18 AATTCAGCTAGAAATCAACAGCTATTTACGTTGGAAATGCGCATGGCTGCG 77
 Db 12 AATATTTGCGGAAACACCTGCGCCAGTCAAGTGGTAATTTACTATGTTGTTAGTGA 71
 QY 78 TCCCATGCGCGTACAGTCAATGACCAATACCGGACAGACGCTGGAACGAGTCA 137
 Db 72 GGAATTTACTACTCAAGCATGACTACTACAAGACACATGATGTGACACCAACAGTAC 131
 QY 138 TCAATCAAGGCGCTGGAAGCGCTTGGCGCTGATATGCTGATCCGTACCGACAT 197
 Db 132 AGAAATTCACCGATTAGAAGAGCGTGTGATGATGTGCGAGTTCTGCTCGTATGA 191
 QY 198 GAGCGCGGCAAGAGCGTTCAACATCAACACAGAGTTACGTTGCGCTGCTGCTGA 257
 Db 192 AGCTGACGCAAGTCTTGTGTCATCAAGAAACAAATCAATATTCGCTGTTGAGGA 251
 QY 258 CATCCACTGACATATGCAATGCGCTGAAGATGAGCGAATACGCGCTCATTTGCTGCG 317
 Db 252 TATTCATTTGATTTACGCAATAGCACTTAAGCATATGATGACGTTTACCAAAATTCG 311
 QY 318 TATTAACCTTGGCAATATGCGTAATGAAGCGTATTCGATGCTGTTGACCTGCGCG 377
 Db 312 GATTAATCTCGTAAACATGATGCTGCGCTGATCGGTTGGAAGATGATTAATCTGCTAA 371
 QY 378 CGATTAACATTCGCAATGCGTATTTGCGTTAACCGCGGATTCCTGGAATAAATCTGCA 437
 Db 372 AGCAAAAAATATTCATTCATTCGTTAATGCTGTTAATGTTAAGAAAAATTAAT 431
 QY 438 AGAAAGATATGCGGAACCGACGCGGCGGTGCTGGAATCTGCATGCGCTATGTA 497
 Db 432 TCAAAATATGTTTACCTTACCTGCTGACGGAATGTAAGAAAGTCACTTCCATATTA 491
 QY 498 TCATCTGATGCGCTGAACCTTCATCAATGTAAGTCAAGCTGAAGCGTTCGATGCT 557
 Db 492 AATTCGCAAGATTTAGATTTTATGATATCATCTTTTGAAGCTTCATGATGA 551
 QY 558 CCTGCTGTTGATCTTATTCGTTTGTGCGCAAAACAGATGATCACCCTGATCTGG 617
 Db 552 TTATGCAATTAAGCTTAATTAAGCTTAATTAATTAATTAATTAATTAATTAATTA 611
 QY 618 GATCACCAGAGCGGTGCTGCGGACGCGGAGTAAATTCGCAATTTGTTAGTCT 677
 Db 612 AATTACAGAACTGTAACCAATTTGCTGAGGAATTAAGAAAGTCTGCTGTTAGAGC 671
 QY 678 GCTGCTGTAAGGATCGGCGACACGCTGCGGATATGCTGCGCGGCGGATCGGTCGA 737
 Db 672 GATACCTAGTTTGGCATGGAATACATTAACAGATATCTTGAAGTCTGATCTGGA 731
 QY 738 AGAGATCAAGTGGTTTGAATTTTGAATGCTGCTATCCGTTGCGAGGATCA 797
 Db 732 AGAAATTAAGTGGCCGCGGAGTTTAAATCAATTTTGTCTTCTCGAATGCCCAAT 791
 QY 798 CTTCATGCGCTGCGGACCGCTGCTGCGGATGATGATGATGATGATGATGATGATGAT 857
 Db 792 GCTTATCTCTGCGCTACTTGGCGGTGAATGAGATTTGATTTAATTCGATGCTAATGA 851

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2003, 15:05:07 ; Search time 1889 Seconds
(without alignments)
9593.833 Million cell updates/sec

Title: US-09-921-992-3
Perfect score: 1119
Sequence: 1 atgcataaccagctcccat.....ttcagcaggttgaaataa 1119

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81.8	7.3	351	14	BU004160 QG637N12.
2	81.6	7.3	720	17	BH235010 PH.05.x.S
3	67.8	6.1	448	10	AV644003 AV644003
4	67.4	6.0	482	10	AV642898 AV642898
5	67.4	6.0	502	10	AV626792 AV626792
6	67.4	6.0	512	10	AV644476 AV644476

7	67.4	6.0	531	10	AV628026
8	67.4	6.0	531	10	AV644087
9	67.4	6.0	545	10	AV626844
10	65.8	5.9	467	10	AV629427
11	62	5.5	377	10	AV641368
12	56.4	5.0	380	9	A1770308
13	54.2	4.8	303	10	AV640429
14	53.8	4.8	597	10	AM282425
15	53.6	4.8	728	12	BG525868
16	53.4	4.8	658	17	A2570993
17	52.6	4.7	640	10	AW738440
18	52.6	4.7	680	12	BG890681
19	51.8	4.6	731	12	BG124857
20	51.4	4.6	492	13	BM323500
21	51	4.6	579	13	B1935552
22	51	4.6	700	13	B1933447
23	50.4	4.5	173	12	BP962770
24	49.8	4.5	576	14	BQ975044
25	49.4	4.4	532	12	BE924238
26	49.4	4.4	645	13	B1934498
27	48.4	4.3	558	10	AV917069
28	48.4	4.3	608	13	B1465443
29	48.4	4.3	616	10	AV934187
30	48.4	4.3	644	13	B1464122
31	48.4	4.3	651	9	AU252386
32	48.4	4.3	672	10	AV916407
33	47.6	4.3	585	10	BE354448
34	47.6	4.3	679	13	B1931456
35	47.4	4.2	534	9	AU252403
36	47.2	4.2	601	10	AM042702
37	46.8	4.2	648	13	B1479404
38	46.6	4.2	670	10	AV821766
39	46.4	4.1	498	10	AM576226
40	45.4	4.1	422	13	B1478402
41	45.4	4.1	528	13	B1482368
42	44.2	3.9	610	13	B1481080
43	44	3.9	465	9	A1482869
44	44	3.9	588	10	AW738232
45	44	3.9	625	10	AW944795

ALIGNMENTS

RESULT 1
BU004160/c 351 bp mRNA linear EST 22-AUG-2002
LOCUS OGG37N12.yg.ab1 QG_EFGHJ Lettuce serriola Lactuca sativa cDNA clone
DEFINITION OGG37N12.yg.ab1 QG_EFGHJ Lettuce serriola Lactuca sativa cDNA clone
ACCESSION BU004160
VERSION BU004160.1 GI:22438555
KEYWORDS EST.
SOURCE Lactuca sativa.
ORGANISM Lactuca sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae; Lactuca.
REFERENCE 1 (bases 1 to 351)
Kozik,A., Michemore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison,P., Koltman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://compgenomics.ucdavis.edu/
Unpublished (2002)
Contact: Alexander Kozik [R.W.Michemore]
Department of Vegetable Crops, R.W.Michemore Lab
University of California at Davis (UCD)
Asmudson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michemore@vegmail.ucdavis.edu]

JOURNAL COMMENT

```
/clone_11b="Chlamydomonas reinhardtii 5%
```

/C:\ONE_110- Citradyomomas terminaluall 26 CO2

QY	50	TTGGGAGTATGCGCGATTGGGAGATGGTGCATCCATCGCCCTATGACATGACCAATAGCC	109
Db	107	TGGGCAAGGTGCGCGGAGGAGCCAGCACCAGCATCGCCCTTACAGCAATGACACACAGG	166
QY	110	GTACGACAGACGCTCGAAGCAGCGTTCATCAATCAAGGCGCTGGAACGCGTTGGCGCTG	169
Db	167	ACACCGCGCAATGTGCGAGCTGACCGCTGACAGGTGACGAAGTGAAGCGCGGCGCGCGG	226
QY	170	ATATCGTCGCTGTATCCGATACCGAGATGAGCGCGGCAACAGCTTCA:-----	217
Db	227	ACATCGTCCGCAATCGACAGTGCAGGGCAGACAGAGGCTGAGGCGTGCATGAAGATTCGGC	286
QY	218	---AATCATCAACAGCAGCGTTAAACGTGCGCTGTGCTGACATTCGACTTCATC	274
Db	287	ACGACGCTGTTCAAGGACAGCATGACAGCTGCCCCCTGATGGCCAGATTCACAGCCCA	346
QY	275	GCATTGCGCTGAAAGTAGCGGGAATACGCGCTGTGATTTCTGCGTATTAACCTGGCATA	334
Db	347	CGGTGGCCCATGTATGTTGGTGGCGG---ACGCGTTGGAAGATTCGTTGAACCCCGGCACT	403
QY	335	TGCGTAATG 343	
Db	404	TGCGTGACG 412	
RESULT 6	AV644476	512 bp	linear
LOCUS	AV644476	512 bp	linear
DEFINITION	AV644476 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii		
ACCESSION	AV644476		
VERSION	AV644476		
KEYWORDS	EST.		
SOURCE	Chlamydomonas reinhardtii.		
ORGANISM	Chlamydomonas reinhardtii.		
REFERENCE	Asanizu, E., Muraoka, K., Inoue, Y., Fukuzawa, H., Ohyama, K., Nakamura, Y. and Tabata, S.		
AUTHORS	Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii		
TITLE	DNA Res. 7 (5), 305-307 (2000)		
JOURNAL	20539644		
MEDLINE	Contact: Erika Asanizu		
COMMENT	The First Laboratory for Plant Gene Research		
FEATURES	Kazusa DNA Research Institute		
source	Yana 1532-3, Kisaizazu, Chiba 292-0812, Japan		
	Email: asanizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/. Location/Qualifiers		
	1. 512		
	/organism="Chlamydomonas reinhardtii"		
	/strain="C9"		
	/db_xref="taxon:3055"		
	/clone="HCL089b10_r"		
	/note="Vector: plasmidscript SK- Site_1: EcoRI; Site_2: XhoI; The cDNA library was constructed from cells cultured in a medium with bubbling air containing 5% carbon dioxide"		
BASE COUNT	118 a 159 g 158 g 77 t		
ORIGIN			
Query Match	6.0%; Score 67.4; DB 10; Length 512;		
Matches 175; Conservative	0; Mismatches 116; Indels 18; Gaps 2;		
QY	50	TTGGGAGTATGCGCGATTGGGAGATGGTGCATCCATCGCCCTATGACATGACCAATAGCC	109
Db	96	TGGGCAAGGTGCGCGGAGGAGCCAGCACCAGCATCGCCCTTACAGCAATGACACACAGG	155
QY	110	GTACGACAGACGCTCGAAGCAGCGTTCATCAATCAAGGCGCTGGAACGCGTTGGCGCTG	169

DB	156	ACACCCGCAATGTGCAAGCTGACCGCTGGACACAGTGAAGAAGTGGCCGGACGCGGGCGCG	215
QY	170	ATATGCTCCGCTATATCCGATACCGACGATGGACCGGACAGCGTTCA-----	217
DB	216	ACATGCTCCGCTACACAGTGCAGGGCCAAAGAGAGAGCTGACGTGATGAATTCGCG	275
QY	218	---AAGTATCAACAGAGCGGTTAAGCGTCCGCTGGTGGCTGACATGCCACTGCGACTATC	274
DB	276	ACGACGCTGTTCCAGACAGCAAGTACGAGTGCCTCTGGTGGCCACATCTCCAGTCCAGCCCA	335
QY	275	GCAATGCGCTGAAAGTACCGGAATACGGCGTGGATTCGTGCTATTAACCTCGCAATA	334
DB	336	CGGTGGCCATGATGTTGGCGG---ACGCGCTTGGAAGAGATCGTGTGAACCCCGCAACT	392
QY	335	TCCGTAATG 343	
DB	393	TTGGTGAAC 401	
RESULT 7			
LOCUS	AV628026	531 bp	mRNA linear EST 15-DEC-2000
DEFINITION	AV628026 Chlamydomonas reinhardtii 5' to 0.04% CO2 Chlamydomonas reinhardtii cDNA clone ICL035f07_r 5', mRNA sequence.		
ACCESSION	AV628026		
KEYWORDS	AV628026.1 GI:10790660		
SOURCE	EST.		
ORGANISM	Chlamydomonas reinhardtii.		
REFERENCE	Chlamydomonas reinhardtii. Eukaryota, Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadaceae; Chlamydomonas.		
AUTHORS	1 (bases 1 to 531) Asamizu,E., Miura,K., Kuchro,K., Inoue,Y., Fukuzawa,H., Ohyama,K., Nakamura,Y., and Tabata,S.		
TITLE	Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii		
JOURNAL MEDLINE	DNA Res. 7 (5), 305-307 (2000)		
COMMENT	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0812, Japan Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/. Location/Qualifiers		
FEATURES	1..531		
Source	/organism="Chlamydomonas reinhardtii" /strain="C9" /db_xref="taxon:3055" /clone="ICL035f07_r" /clone_lib="Chlamydomonas reinhardtii 5' to 0.04% CO2" /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI; The cDNA library was constructed from cells cultured in a carbon stress acclimatized condition in which carbon dioxide concentration in the bubbling gas was changed from 5% to 0.04%"		
BASE COUNT	119 a	166 c	165 g 81 t
ORIGIN			
Query Match	6.0%; Score 67.4; DB 10; Length 531;		
Best Local Similarity	56.6%; Pred. No. 1.8e-09;		
Matches 175; Conservative 0; Mismatches 116; Indels 18; Gaps 2;			
QY	50	TTGGGAATGTCGATGGCGATGGGTGTCCTCATCGCCGCTGACATCATGCAATACGC	109
DB	106	TTGGCAAGCTCCGGTGGGACGACGACCGCATCGCCCTTCAACATGACCAACACGCG	165
QY	110	GTAGACAGAGCTGGAAGCAACGGTCAATCAATCAAGAGCGCTGACAGCGCTGGCGCTG	169
DB	166	ACACCCGGAATGTCAGCTGACCTGGACCAAGGTGAAGAAATGCGCGGACCGCGCGCG	225
QY	170	ATATGCTCCGCTATATCCGATACCGACGATGGACGCGGCAAGAGCGTTCA-----	217
DB	226	ACATGCTCCGCTACACAGTGCAGGGCCAAAGAGAGGCTGAGGCGTCATGAAGATTGCGG	285

QY	218	---	AAACATCAAAACAGAGGTTAACGTCGGCGGCGGCGGCGGATCATCATCAGCATATC	274
Db	286	AGCAGCTGTTTAAAGACAGACATGACGACGGCCCGTGGTGCCGACATTCACATCCAGGCCA	345	
QY	275	GCATTTGGCGCTGAAGATGAGCGGAAATACGGCGTCGATTGCTGCTGATTAAACCTGGCAATA	334	
Db	346	CGGTGGCGCATGATGATGGTGGCGG---	ACGGCTTCGAGAAGATCCGTGTGAACCCCGGCAACT	402
QY	335	TCGGTAATG	343	
Db	403	TCGCTGACG	411	

RESULT	8
LOCUS	AV644087
DEFINITION	AV644087 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
ACCESSION	AV644087
VERSION	AV644087
KEYWORDS	AV644087.1 GI:10787415
SOURCE	EST.
ORGANISM	Chlamydomonas reinhardtii.
REFERENCE	Chlamydomonas reinhardtii.
AUTHORS	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
	Chlamydomonadaceae; Chlamydomonas.
	1 (bases 1 to 531)
	Asanizu, E., Miura, K., Kuchio, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,
	Nakamura, Y. and Tabata, S.
	Generation of expressed sequence tags from low-CO2 and high-CO2
	adapted cells of Chlamydomonas reinhardtii
	DNA Res. 7 (5), 305-307 (2000)
TITLE	20539644
JOURNAL	Contact: Erika Asanizu
MEDLINE	The First Laboratory for Plant Gene Research
COMMENT	Kazusa DNA Research Institute
	Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
	Email: asanizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES	Location/Qualifiers
source	1..531
	/organism="Chlamydomonas reinhardtii"
	/strain="C9"
	/db_xref="taxon:3055"
	/clone="HCL081e11_r"
	/clone_11b="Chlamydomonas reinhardtii 5% CO2"
	/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
	XhoI. The cDNA library was constructed from cells cultured
	in a medium with bubbling air containing 5% carbon
	dioxide"

BASE COUNT	120 a	168 c	162 g	81 t
ORIGIN				
Query Match				
Best Local Similarity	56.6%			
Matches 175; Conservative	0;			
Mismatches 116;				
Indels 18;				
Gaps 2;				
QY	50	TTGGGAATGTGCGCATTTGGCGATGTGCTGCCATGCGCGTACAGTCAATGACCAATACGC	109	
Db	97	TTCGGCAAGTGGCGGTGGGACACGACACCGGATGCGCTTCACACATGACACACACGAG	156	
QY	110	GTACGACAGAGCTGCAACCAACGGTCAATCAATTAAGGCGCTGGAACGCGTTGGCGGTG	169	
Db	157	ACACCCGCAATGTGACGCTGACCCGTGGACACAGGTGTAAACAAGTCCCGGACCGGCGCG	216	
QY	170	ATATGTCGCGGTATTCGTTACCGAGAGTGGAGCGGCGGAGAAGCGTTCA-----	217	
Db	217	ACATGTCGCGCATTCACAGTGCAGGCGCAAGAAAGAGGCTGAGGCGGTGCATTGAAGATTCCG	276	
QY	218	---AACTCATCAAAACAGAGGTTAACGTCGGCGGCGGCGGCTGACATTCACATTCGACTATC	274	
Db	277	AGCAGCTGTTTAAAGACAGACATGACGACGGCCCGTGGTGCCGACATTCACATCCAGGCCA	336	
QY	275	GCATTTGGCGCTGAAGATGAGCGGAAATACGGCGTCGATTGCTGCTGATTAAACCTGGCAATA	334	

Db	337	CGGTCGCAATGATGCTGGCGS---	ACGCCTTCGAGAAAGATCCGTGTGAACCCCGCAACT	393
QY	335	TCGCTAATG 343		
Db	394	TCGCTGACG 402		
RESULT 9				
LOCUS	AV626844			
DEFINITION	AV626844 Chlamydomonas reinhardtii 5% to 0.04% CO2 Chlamydomonas reinhardtii cDNA clone lcl016e04_r 5', mRNA sequence.			
ACCESSION	AV626844			
VERSION	AV626844.1	GI:10789124		
KEYWORDS	EST.			
SOURCE	Chlamydomonas reinhardtii.			
ORGANISM	Chlamydomonas reinhardtii			
REFERENCE	Eukaryote: Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadaceae; Chlamydomonas.			
AUTHORS	1 (bases 1 to 545) Asamizu, E., Miura, K., Kuchio, K., Inoue, Y., Fukuzawa, H., Ohyama, K., Nakamura, Y. and Tabata, S.			
TITLE	Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii			
JOURNAL	DNA Res. 7 (5), 305-307 (2000)			
MEDLINE	20539644			
COMMENT	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yama 153-3, Kisarazu, Chiba 292-0812, Japan Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/. Location/Qualifiers			
FEATURES				
Source	1..545 /organism="Chlamydomonas reinhardtii" /strain="C9" /db_xref="taxon:3055" /clone="lcl016e04_r" /clone_lib="Chlamydomonas reinhardtii 5% to 0.04% CO2" /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI; The cDNA library was constructed from cells cultured in a carbon stress acclimatized condition in which carbon dioxide concentration in the bubbling gas was changed from 5% to 0.04%"			
BASE COUNT	121 a 177 c 166 g 81 t			
ORIGIN				
Query Match	6.0%: Score 67.4; DB 10; Length 545;			
Best Local Similarity	56.6%: Pred. No. 1.8e-09;			
Matches 175: Conservative	0; Mismatches 116; Indels 18; Gaps 2;			
QY	50	TTGGGATGATGCCGATTTGGCGATGCTGCCATCGCGTACATGCATGACCAATACGC	109	
Db	90	TCGGCAAGTGTGGCGGTGGCGACGACGACCCGATCGCCTTCAGACATGACACACGCG	149	
QY	110	GTACGACAGAGCTGACAGCAACGCGTCATCAATTAAGCCCTGGAACGCGTGGCGCTG	169	
Db	150	ACACCCGGAATGTGAGCTGACCGGTGGACAGGTAAAGTGGCCGGACCCGGCGCGG	209	
QY	170	ATATGTCGCGTATCCGTACCGAGACGATGGAGCGCGGAGAAAGCGTTCA-----	217	
Db	210	ACATGCTCCGATCCATCAATGACGAGGCAAGAAAGAGGCTGAGAGCGTGCATGAAGATTCCG	269	
QY	218	---AATCATCAAAACAGCAGGTTAAAGTGGCGCGTGGCTGACATCCACTTCGACTATC	274	
Db	270	AGCACTCTTCAAGAGCAAGTACAGACGTCGCGCTGTGGCGACATCCACTTCAGGCCCA	329	
QY	275	GCATTGGCGCTGAAGTACCGGAATAGCGCGTGCATGTTCTGCTGATTAAACCTGGCAATA	334	
Db	330	CGGTGGCGCATATGTTGGCGG---ACGCTTCGAGAAAGATCCGTGTGAACCCCGCAACT	386	
QY	335	TCGCTAATG 343		

RESULT 10	AV629427	467 bp	mRNA	linear	EST 15-DEC-2000
LOCUS	AV629427				
DEFINITION	AV629427 Chlamydomonas reinhardtii	58 to 0.04% CO2	Chlamydomonas		
ACCESSION	reinhartd11.cdna clone LCL058f01_r 5'		mRNA sequence.		
VERSION	AV629427				
KEYWORDS	AV629427.1	GI:10792061			
SOURCE	EST.				
ORGANISM	Chlamydomonas reinhardtii.				
REFERENCE	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;				
AUTHORS	Chlamydomonas reinhardtii.				
TITLE	1 (bases 1 to 467)				
JOURNAL	Asamizu, E., Miura, K., Kuchio, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,				
MEDLINE	Nakamura, Y. and Tabata, S.				
COMMENT	Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii				
FEATURES	DNA Res. 7 (5), 305-307 (2000)				
SOURCE	Contact: Erika Asamizu				
	The First Laboratory for Plant Gene Research				
	Kazusa DNA Research Institute				
	Yama 1532-3, Kisarazu, Chiba 292-0812, Japan				
	Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.				
	Location/Qualifiers				
	1. 467				
	/organism="Chlamydomonas reinhardtii"				
	/strain="C9"				
	/db_xref="taxon:3055"				
	/clone="LCL058f01_r"				
	/note="Vector: pBluescriptII SK-; Site1: EcoRI; Site2: XhoI; The cDNA library was constructed from cells cultured in a carbon stress acclimatized condition in which carbon dioxide concentration in the bubbling gas was changed from 5% to 0.04%"				
BASE COUNT	108 a 149 c 142 g 68 t				
ORIGIN					
Query Match	5.9% Score 65.8; DB 10; Length 467;				
Best Local Similarity	56.3%; Pred. No. 5.1e-09;				
Matches 174; Conservative	0; Mismatches 117; Indels 18; Gaps 2;				
QY	50 TTGGCAATGTGCGCATGTCGTCCTCCATCGCCGTCAGTCATGCAACAATACG 109				
Db	93 TCGGCAAGTGGCGGTGGGACGACGACCGCATGCGCCCTTCAAGACATGACACACG 152				
QY	110 GTACGACAGAGCTGCAACACGTCATCAATCAATCAAGGCGCTTGGCGCTG 169				
Db	153 ACACCCGGAATGTGAGTGCAGTCGCGGACCGAGTGAAACATGTGGCGGCGCGG 212				
QY	170 AATATGTCCTGTATCCCTTACCGAGATGGAGCGGCGAGAACGCTTCA----- 217				
Db	213 ACATGTCCTGCATCATCATGTCAGGCAAGAGAGCGTCAGAGCGTCATGAAGATTGCG 272				
QY	218 ---AATCATCAAAACAGCAGGTTAAAGCTGCGCTGGTGGCTGACATCCACTTGCATTATC 274				
Db	273 AGCAGCTGTTCAAGGACAAATGACGACGTCGTCGTCGTCGATCCACTT---CCAGC 329				
QY	275 GCATTGGCCTCAAGTACGGAATACGGCGTCGATGTCCTGCGATTAAACCTTGGCATA 334				
Db	330 CCACGCGCGCATGATGTTGGCGGACGCTTCGAGAAAGATCGTGTGAACCCCGCAACT 389				
QY	335 TCGGTAATG 343				
Db	390 TCGGTGACG 398				
RESULT 11					
	AV641368				

LOCUS	AV641368	377 bp	mRNA	linear	EST 15-DEC-2000
DEFINITION	AV641368 Chlamydomonas reinhardtii 5' CO2 cDNA clone HCL032d09_r 5', mRNA sequence.				
ACCESSION	AV641368				
VERSION	AV641368.1	GI:10784696			
KEYWORDS	EST.				
SOURCE	Chlamydomonas reinhardtii.				
ORGANISM	Chlamydomonas reinhardtii.				
REFERENCE	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadaceae; Chlamydomonas.				
AUTHORS	1 (bases 1 to 377)				
TITLE	Asamizu,E., Miura,K., Kuchio,K., Inoue,Y., Fukuzawa,H., Ohyama,K., Nakamura,T., and Tabeta,S.				
JOURNAL	Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii				
MEDLINE	DNA Res. 7 (5), 305-307 (2000)				
COMMENT	20539644				
FEATURES	Contact: Erika Asamizu				
SOURCE	The First Laboratory for Plant Gene Research				
	Kazusa DNA Research Institute				
	Yama 1532-3, Kisarazu, Chiba 292-0812, Japan				
	Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.				
	Location/Qualifiers				
	1..377				
	/organism="Chlamydomonas reinhardtii"				
	/strain="C9"				
	/db_xref="taxon:3055"				
	/clone="HCL032d09_r"				
	/clone_lib="Chlamydomonas reinhardtii 5% CO2"				
	/note="vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI; The cDNA library was constructed from cells cultured in a medium with bubbling air containing 5% carbon dioxide"				
BASE COUNT	89 a	118 c	116 g	54 t	
ORIGIN					
Query Match	5.5%;	Score 62;	DB 10;	Length 377;	
Best Local Similarity	55.6%;	Pred. No. 6;	e-08;		
Matches 150;	Conservative	0;	Mismatches 105;	Indels 15;	Gaps 1;
OY	50 TTGGCAATGTGGCGGCGTGGTGGCGATGTGCTGCCATCGCGCGTACAGTCCATGACATGACCAATACGC	109			
DB	102 TCGGGCAAGGTGGCGGTGGGACGCCACGACGCCCATCGCCCTTACAGACATGACACACACGCG	161			
OY	110 GTACGACAGACAGCTGCACCAACAGCGTCAATCAATCAAGCGCGCTGGAAACGCGTTGGCGCTG	169			
DB	162 ACACCCGGCAAGTGTGAGCTGACCGCTGGACACGAGTGAAGAAATATGGCGGACCGCGCGGG	221			
OY	170 ATATGTCGCGGTATCGCTACCGACGATGACGCGGCGCAAGACGCTCA-----	217			
DB	222 ACATGTCGCGCATCACTGATCGAGGCAAGAGGAGGCTGACGCGTGCATGAACATTGCGC	281			
OY	218 ---AACTCATCAACAGAGAGGTTAACGCGCGCGTGGTGGCGATCGACATCGACTGACATC	274			
DB	282 AGCAGCTCTTCAAGGACAGTACGACGCGCCCTGGTGGCGACGACATCCATCCAGGCCA	341			
OY	275 GCATTGCGCGTGAAGTAGACGAATACGCG	304			
DB	342 CGGTGGCATGATGTGGCGGACGCGCTTCG	371			
RESULT 12					
LOCUS	AI1770308	380 bp	mRNA	linear	EST 24-JAN-2000
DEFINITION	35 Mycobacterium anaerobic stationary phase library Mycobacterium smegmatis cDNA, mRNA sequence.				
ACCESSION	AI1770308				
VERSION	AI1770308.1	GI:6742677			
KEYWORDS	EST.				
SOURCE	Mycobacterium smegmatis.				
ORGANISM	Mycobacterium smegmatis				
	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.				

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2003, 15:09:17 ; Search time 78 Seconds
(without alignments)
4399.632 Million cell updates/sec

Title: US-09-921-992-3
Perfect score: 1119
Sequence: 1 atgcataaccaggtcccaat.....ttcagcaggttgaaataa 1119

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCBUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1010	90.3	1010	2	US-08-827-190-8
2	1010	90.3	1010	4	US-09-170-187-8
3	631.6	56.4	1345	2	US-08-827-190-4
4	631.6	56.4	1345	4	US-08-827-190-4
5	619.8	55.4	1079	2	US-08-827-190-7
6	619.8	55.4	1079	4	US-08-827-190-7
7	534.6	47.8	886	2	US-08-827-190-10
8	534.6	47.8	886	4	US-08-827-190-10
9	269	24.0	886	2	US-08-827-190-9
10	269	24.0	886	4	US-08-827-190-9
11	55.4	5.0	1083	4	US-09-170-187-9
12	40.6	3.6	1206	4	US-09-221-017B-309
13	40.6	3.6	2299	4	US-09-153-589A-2
14	38.4	3.4	1665	4	US-09-153-589A-1
15	38.2	3.4	4411529	4	US-09-221-017B-1017
16	35.4	3.2	583	4	US-09-103-840A-1
17	35.4	3.2	737	4	US-08-998-416-1089
18	35	3.1	1383	3	US-08-998-416-1039
19	35	3.1	1383	3	US-09-400-742-7
20	35	3.1	1383	4	US-08-618-651A-7
21	34	3.0	8910	3	US-08-215-252-6
22	34	3.0	4411529	4	US-08-779-764A-1
23	33.8	3.0	44377	2	US-08-804-227C-7
24	33.8	3.0	44377	2	US-08-804-227C-7
25	33.6	3.0	606	2	US-08-432-871C-48
26	33.6	3.0	606	2	US-08-270-956-48
27	33.4	3.0	1660	4	US-09-297-468-1

C 28	33.4	3.0	23673	4	US-09-773-816-1	Sequence 1, Appli
C 29	32.8	2.9	71989	4	US-09-443-501A-2	Sequence 2, Appli
C 30	32.4	2.9	51335	3	US-08-369-822C-9	Sequence 9, Appli
C 31	32.4	2.9	51335	3	US-08-582-776C-9	Sequence 9, Appli
C 32	32.4	2.9	51335	3	US-08-434-818-9	Sequence 9, Appli
C 33	32.4	2.9	5698	1	US-08-761-258-11	Sequence 11, Appli
C 34	32.4	2.9	5698	2	US-08-977-306-11	Sequence 11, Appli
C 35	32.4	2.9	8910	3	US-08-369-822C-19	Sequence 19, Appli
C 36	32.4	2.9	8910	3	US-08-582-776C-19	Sequence 19, Appli
C 37	32.4	2.9	8910	3	US-08-434-818-19	Sequence 19, Appli
C 38	32.4	2.9	4403765	4	US-09-103-840A-2	Sequence 2, Appli
C 39	31.8	2.8	4403765	4	US-09-103-840A-2	Sequence 2, Appli
C 40	31.6	2.8	43280	2	US-08-804-227C-1	Sequence 1, Appli
C 41	31.4	2.8	1912	1	US-08-270-013B-1	Sequence 1, Appli
C 42	31.4	2.8	1912	1	US-08-838-418-1	Sequence 1, Appli
C 43	31.2	2.8	1778	5	PCT-US91-02958-9	Sequence 9, Appli
C 44	31.2	2.8	3934	3	US-09-226-568-18	Sequence 18, Appli
C 45	31.2	2.8	3946	1	US-08-077-848A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-827-190-8
Sequence 8, Application US/08827190
Patent No. 5858367
GENERAL INFORMATION:
APPLICANT: Rather, Phillip N.
TITLE OF INVENTION: Methods For Screening For Antimicrobials
TITLE OF INVENTION: Utilizing aarc And Compositions Thereof
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/827,190
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: CASE-02443
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1010 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-827-190-8
Query Match 90.3% Score 1010; DB 2; Length 1010;
Best Local Similarity 100.0%; Pred. No. 5.8e-309; Indels 0; Gaps 0;
Matches 1010; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ATGCATACCAAGGCTCAATTCAGTAGAAGAAATCAACACGTAATTCAGTGGATGTG 60
Db 1 ATGCATACCAAGGCTCAATTCAGTAGAAGAAATCAACACGTAATTCAGTGGATGTG 60
OY 61 CGAATGGCATGTGTCTCCATCGCCGTACAGTCCATGACCAATATACGATACGACAGAC 120

|||||
Db 61 CCGATTGGCGATGGTGTCTCCATCGCCGTACAGTCCATGACCAATACGCGTACAGACAGAC 120
QY 121 GTGGAACCAACGGTGCATCAATCAACAGCCGCTGGAACCGCTTGGCGCTGATATCGCT 180
Db 121 GTGGAACCAACGGTGCATCAATCAACAGCCGCTGGAACCGCTTGGCGCTGATATCGCT 180
QY 181 GTATCCGTACCGACGATGAGAGCGGCGAGAACGCTTCAACTCATCAACAGCAGGTTAAC 240
Db 181 GTATCCGTACCGACGATGAGAGCGGCGAGAACGCTTCAACTCATCAACAGCAGGTTAAC 240
QY 241 GTGCGCGTGTGGCTGACATCCACTTGCATTCGACTATTCGCTGCTGAAAGTACCGGAATAC 300
Db 241 GTGCGCGTGTGGCTGACATCCACTTGCATTCGACTATTCGCTGCTGAAAGTACCGGAATAC 300
QY 301 GGGGTGATGTGTGGCTGATTAACCCGGAATATGAGTATGAGAGCGCTATTCGATG 360
Db 301 GGGGTGATGTGTGGCTGATTAACCCGGAATATGAGTATGAGAGCGCTATTCGATG 360
QY 361 GTGATTGACTGTGCGCGCGATTAACCAATTCGATCCGCTTGGCTTAAACGCGGATTCG 420
Db 361 GTGATTGACTGTGCGCGCGATTAACCAATTCGATCCGCTTGGCTTAAACGCGGATTCG 420
QY 421 CTGGAAGAAAGATCTGCAAGAAAGTATGGCAACCGACGCGGCGGCTTGTGGAATCT 480
Db 421 CTGGAAGAAAGATCTGCAAGAAAGTATGGCAACCGACGCGGCGGCTTGTGGAATCT 480
QY 481 GCCATGGGTATGTGATCTGATCTGATCGATCGCTGATGATGATGATGATGATGATGATG 540
Db 481 GCCATGGGTATGTGATCTGATCTGATCGATCGCTGATGATGATGATGATGATGATGATG 540
QY 541 AAAGCGTCTAGCGTCTCTCCGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 600
Db 541 AAAGCGTCTAGCGTCTCTCCGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 600
QY 601 CACCGCTGATCTGGGGATCACCGAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 601 CACCGCTGATCTGGGGATCACCGAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 661 GCCATTGCTTAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 661 GCCATTGCTTAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 721 GCGGCGGATCCGCTCGAAGAGATCAAAAGTGGTTTGAATTTTGAATTTTGAATTTTGAATTT 780
Db 721 GCGGCGGATCCGCTCGAAGAGATCAAAAGTGGTTTGAATTTTGAATTTTGAATTTTGAATTT 780
QY 781 CGTTCGAGAGGATCAACTCATCGCTGCGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 781 CGTTCGAGAGGATCAACTCATCGCTGCGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 841 ATCGTACCGTAAACGCGCTGAGAGCAACGCGCTGAGAGATATCATCACTCGATGAGAGT 900
Db 841 ATCGTACCGTAAACGCGCTGAGAGCAACGCGCTGAGAGATATCATCACTCGATGAGAGT 900
QY 901 TCGATTATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 TCGATTATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 961 ACCGGGGGCAACAAAGAAAGCGGCTCTATGAGATGCGCTGGGCAAGA 1010
Db 961 ACCGGGGGCAACAAAGAAAGCGGCTCTATGAGATGCGCTGGGCAAGA 1010

RESULT 2

US-09-170-187-8
; Sequence 8, Application US/09170187
; Patent No. 6383745
; GENERAL INFORMATION:
; APPLICANT: Rather, Philip N.
; TITLE OF INVENTION: Methods For Screening For Antimicrobials
; TITLE OF INVENTION: Utilizing aarc And Compositions Thereof
; NUMBER OF SEQUENCES: 13

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Medlen & Carroll, LLP
;; STREET: 220 Montgomery Street, Suite 2200
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: United States of America
;; ZIP: 94104
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: IBM PC compatible
;; SOFTWARE: Patent Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/170,187
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/827,190
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Carroll, Peter G.
;; REGISTRATION NUMBER: 32,837
;; REFERENCE/DOCKET NUMBER: CASE-02443
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 705-8410
;; TELEFAX: (415) 397-8338
;; INFORMATION FOR SEO ID NO: 8
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1010 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: unknown
;; MOLECULE TYPE: DNA (genomic)
;; US-09-170-187-8

Query Match 90.3%; Score 1010; DB 4; Length 1010;
Best Local Similarity 100.0%; Pred. No. 5,8e-309;
Matches 1010; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATATACAGGAGTCCCAATTCACAGTAAATCAACAGCTATTTACGTTGGGAATGTG 60
Db 1 ATGATATACAGGAGTCCCAATTCACAGTAAATCAACAGCTATTTACGTTGGGAATGTG 60
QY 61 CCGATTGGCGATGGTGTCTCCATCGCCGTACAGTCCATGACCAATATGCGCTACGACAGAC 120
Db 61 CCGATTGGCGATGGTGTCTCCATCGCCGTACAGTCCATGACCAATATGCGCTACGACAGAC 120
QY 121 GTGGAACCAACGGTGCATCAATCAACAGCCGCTGGAACCGCTTGGCGCTGATATCGCT 180
Db 121 GTGGAACCAACGGTGCATCAATCAACAGCCGCTGGAACCGCTTGGCGCTGATATCGCT 180
QY 181 GTATCCGTACCGACGATGAGAGCGGCGAGAACGCTTCAACTCATCAACAGCAGGTTAAC 240
Db 181 GTATCCGTACCGACGATGAGAGCGGCGAGAACGCTTCAACTCATCAACAGCAGGTTAAC 240
QY 241 GTGCGCGTGTGGCTGACATCCACTTGCATTCGACTATTCGCTGCTGAAAGTACCGGAATAC 300
Db 241 GTGCGCGTGTGGCTGACATCCACTTGCATTCGACTATTCGCTGCTGAAAGTACCGGAATAC 300
QY 301 GGGGTGATGTGTGGCTGATTAACCCGGAATATGAGTATGAGAGCGCTATTCGATG 360
Db 301 GGGGTGATGTGTGGCTGATTAACCCGGAATATGAGTATGAGAGCGCTATTCGATG 360
QY 361 GTGATTGACTGTGCGCGCGATTAACCAATTCGATCCGCTTGGCTTAAACGCGGATTCG 420
Db 361 GTGATTGACTGTGCGCGCGATTAACCAATTCGATCCGCTTGGCTTAAACGCGGATTCG 420
QY 421 CTGGAAGAAAGATCTGCAAGAAAGTATGGCAACCGACGCGGCGGCTTGTGGAATCT 480
Db 421 CTGGAAGAAAGATCTGCAAGAAAGTATGGCAACCGACGCGGCGGCTTGTGGAATCT 480
QY 481 GCCATGGGTATGTGATCTGATCTGATCGATCGCTGATGATGATGATGATGATGATGATG 540
Db 481 GCCATGGGTATGTGATCTGATCTGATCGATCGCTGATGATGATGATGATGATGATGATG 540

```

Db 481 GCCATGCGTATGTTGATCATCTGCATCGCTGAACTTCGATCAGTCAAGTCAGCGTG 540
Oy 541 AAAGCGTCTGACGCTCTTCCTGCTGTTGATGCTTATCGTTTCTGGCAAAACAGATCGAT 600
Db 541 AAAGCGTCTGACGCTCTTCCTGCTGTTGATGCTTATCGTTTCTGGCAAAACAGATCGAT 600
Oy 601 CAGCGCTTGACATCTGGGGATCACCGAAGCGGTGTGCGCGAGCGGGGAGTAAATCC 660
Db 601 CAGCGCTTGACATCTGGGGATCACCGAAGCGGTGTGCGCGAGCGGGGAGTAAATCC 660
Oy 661 GCCATTGTTGTTAGTCTGCTGCTGTCTGAAGCATCGCGACACGCTGCCGTATCGCTG 720
Db 661 GCCATTGTTGTTAGTCTGCTGCTGTCTGAAGCATCGCGACACGCTGCCGTATCGCTG 720
Oy 721 GCGGCGCATCCGCTGGAAGAGATCAAAAGTGGTTTCGATATTTTGAATCGCTGCCTATC 780
Db 721 GCGGCGCATCCGCTGGAAGAGATCAAAAGTGGTTTCGATATTTTGAATCGCTGCCTATC 780
Oy 781 CGTTCGCGAGGATCAACTTCATCGCTGCGCGACCTGTTGCGCTCAGAGATTTGATGTT 840
Db 781 CGTTCGCGAGGATCAACTTCATCGCTGCGCGACCTGTTGCGCTCAGAGATTTGATGTT 840
Oy 841 ATCGGTACGCTTAACGCGCTGAGACAAAGCCCTGGAAGATATCATCTCCGATGGACGTT 900
Db 841 ATCGGTACGCTTAACGCGCTGAGACAAAGCCCTGGAAGATATCATCTCCGATGGACGTT 900
Oy 901 TCGATTATGCGCTGCTGCTGATGAGCCAGGTGAGAGCGCTGTTCTACACTCGGCGCTC 960
Db 901 TCGATTATGCGCTGCTGCTGATGAGCCAGGTGAGAGCGCTGTTCTACACTCGGCGCTC 960
Oy 961 ACCGCGCGCAACAGAAAAGCGGCTCTATGAAGATGCGCGCAAGA 1010
Db 961 ACCGCGCGCAACAGAAAAGCGGCTCTATGAAGATGCGCGCAAGA 1010

```

RESULT 3
US-08-827-190-4
Sequence 4, Application US/08827190
Patent No. 5858367

GENERAL INFORMATION:

APPLICANT: Rather, Philip N.
TITLE OF INVENTION: Methods For Screening For Antimicrobials
TITLE OF INVENTION: Utilizing aarc And Compositions Thereof
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:

ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco

STATE: California
COUNTRY: United States of America
ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/827,190
FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: CASE-02443

TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1345 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 175..1272
OTHER INFORMATION: /gene="aarc"
OTHER INFORMATION: /note="Similar to E. coli Gpe protein listed by GenBank."
US-08-827-190-4

Query Match
Best Local Similarity 56.4%; Score 631.6; DB 2; Length 1345;
Matches 809; Conservative 0; Mismatches 274; Indels 3; Gaps 1;

```

Oy 1 ATGCATACCAAGGCTCCATTCATACGTAGAAATCAACAGTATTACGTGGGATGTG 60
Db 175 ATGCATATGATGATCCAGCTTAAAGAGCTAAATCCAGCAATTTATAGTAAACGTG 234
Oy 61 CCGATTGGGATGAGTGTCTCCATCGCGGTACAGTCCATACCAATACGGGTACGACAGC 120
Db 235 CCTATTGGGATGAGTGTCTCCATCGGTACAGTCCATACCAATACGAGCAAGAGAT 294
Oy 121 GTGAGCAACAGGCTCCATCAATCAAGGCGTGAAGCGGCTGATATCGCTCGT 180
Db 295 GTTGAAGCCACTGTGTGGGCAATTCATCTTAGACGTGTAGTGTATATCGTCCGC 354
Oy 181 GTATCCGTACCGACGATGAGCGCGGAGAGGCTTCAACATCATCAACAGCAGTTAAC 240
Db 355 GTGTCTGTTCCTACGATGATGATGAGCAGAGAGCCTTAAATTAATTAAGCAGCGCTGAA 414
Oy 241 GTGCGCTGTGTGTGATCATCTTCATCTGATCGATTTGCGGTGAAGTACGGAATAC 300
Db 415 GTGCCATTGTGTGCGGATTTCACTTTGACTACCGTATCGCATGAACTGTGCTGAATAT 474
Oy 301 GCGGTGATTTGTGTGCTGATTAACCTGGCAATATGTAAGAGGATATTCGATG 360
Db 475 GGTGTGACTGCTTACGAATTAACCGAGTAATATGCGAGTGAAGAGCTATTCGCCAA 534
Oy 361 GTGTGACTGTGCGCGCATAAACATTCGATCCGATTATGCGCTTAACCGCGGATG 420
Db 535 GTGTGATAGTGTGTGCTGATCACAACATTCGATATGAGGGTTCATATGGCGGATCA 594
Oy 421 CTGGAAGAAATCTGCAAGAAAGTATGGCGAAGCCAGCGCGGCTGCTGCAATCT 480
Db 595 CTGGAAGAAATCTGCAAGAAAGTATGGCGAAGCCAGCGCGGCTGCTGCAATCT 654
Oy 481 GCCATGCGTATGATCATCTGATGCGGCAATTTGATGATCAAGTCAACGACGCTG 540
Db 655 GCATGCGATGATGATCTTGTGACAGCTGAATTTGATGATCAAGTCAACGACGCTG 714
Oy 541 AAAGCTCTGAGCTTCTCCTGCTGTTGATGCTTATCGTTTGTGCAAAACAGATCGAT 600
Db 715 AAAGCTCTGAGCTTCTCCTGCTGTTGATGCTTATCGTTTGTGCAAAACAGATCGAT 774
Oy 601 CAGCGCTGATCTGGGGATCACCGAAGCGGTGTGCGCGACGCGGCGAGTAAATCC 660
Db 775 CAACCACTTCACCTCGTATTCAGAAAGCGGGGCTGCTGCTGCTGCAAGAAATCA 834
Oy 661 GCCATTGTTGTTAGTCTGCTGCTGTGTAAGAGTGGGAGCAACGCTGGCGATGCTG 720
Db 835 GCATTTGCTTGTGATGTTGTTGCTGTAAGATGCGGATACGTTACGTATCTACTC 894
Oy 721 GCGGCGCATCCGCTGGAAGAGATCAAAAGTGGTTTCGATATTTTGAATCGCTGCCTATC 780
Db 895 GCGGCGCATCCGCTGGAAGAGATCAAAAGTGGTTTCGATATTTTGAATCGCTGCCTATC 954
Oy 781 CGTTCGCGAGGATCAACTTCATCGCTGCGGACCTGTTGCGCTGAGAAATTTGATGTT 840
Db 955 CGCTACGCGGATCAACTTCATCGCTGCGGACCTGTTGCGCTGAGAAATTTGATGTTG 1014
Oy 841 ATCGGTACGCTTAACGCGCTGAGACAAAGCTTGAAGATATCATCTCCGATGGACGTT 900
Db 1015 ATTTGATCGGTAAATCTTTGGAGACGCGCTTGAAGATATTTATCATCGCGGATGATGTC 1074
Oy 901 TCGATTATGCGCTGCTGCTGATGAGCCAGGTGAGGCGCTGTTCTACACTCGGCGCTC 960

```

Db 1075 TCTATTATTTGTTGTGTAAGGAATGGCCGCGGTGAAGCCGAGGTTTCTACTTTAGGCTG 1134
QY 961 ACCGGCGGCAACAAGAGGGGCTCTATGAAGATGCGGTGGC---AAAGACCGCTG 1017
Db 1135 GCTGGCGCGAAACCAAAAGGTTTCTATGAGATGCGCTTCCCAAAAAGAGCGTTT 1194
QY 1018 GACAAACAGATATGATGCGAGCTGGGAAGCAGCATTCGTGGCAAGCCAGTCAGCTG 1077
Db 1195 GATAATGACATATTTATTTGATCAGCTTGAGCGCAAAATTCGGCGAAAGCAGCAATGCTT 1254
QY 1078 GACGAA 1083
Db 1255 GATGAA 1260

RESULT 4
US-09-170-187-4
Sequence 4, Application US/09170187
Patent No. 6383745

GENERAL INFORMATION:
APPLICANT: Rafter, Philip N.
TITLE OF INVENTION: Methods For Screening For Antimicrobials
TITLE OF INVENTION: Utilizing aarc And Compositions thereof
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/09/170,187
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/827,190
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: CASE-02443
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-0410
TELEFAX: (415) 397-8398

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1345 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:

NAME/KEY: CDS
LOCATION: 175..1272
OTHER INFORMATION: /gene="aarc"
OTHER INFORMATION: /note="Similar to E. coli Gcpe protein listed by GenBank."
US-09-170-187-4

Query Match 56.43; Score 631.6; DB 4; Length 1345;
Best Local Similarity 74.53; Pred. No. 2e-189;
Matches 809; Conservative 0; Mismatches 274; Indels 3; Gaps 1;

QY 1 ATGCATACAGGCTCCATTCATGAGAAATCAACAGCATTTTACGTTGGAAATGTG 60
Db 175 ATGCATATATGATCAGGCAAAAGAGAGCTAATCCACCGAATTTATGATGAGTGTG 234

QY 61 CCGATTGGCGATGCTGCTCCATTCGCGTACAGTCAATGACCAATACCGGTACGACAG 120
Db 235 CTTATTTGGATGCTGCTCCATTCGCTCCATTCATCTATGAGAAATACGCGACAGCGAT 294
QY 121 GTCAAGCAACGCTCATCAATCAAGGCGCTGGAAGCGCTGGCGCTGATGCTGCT 180
Db 295 GTTGAAGCAACGCTGCGGCAATCAATCAATCAATGAGGCTGATGCTGCTGCTGCT 354
QY 181 GTATCCGTACGAGATGAGCGGCGGAGAGCGCTTCAATCAATCAATCAATCAAT 240
Db 355 GTCTCTTCTCTACGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 414
QY 241 GTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 415 GTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 474
QY 301 GCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 475 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 534
QY 361 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 535 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 594
QY 421 CTGCAAAAGATCTGCAAGAAAGATGCGAAGCGGCGGCGGCGGCTGCTGCTGCT 480
Db 595 CTGCAAAAGATCTGCAAGAAAGATGCGAAGCGGCGGCGGCGGCTGCTGCTGCT 654
QY 481 GCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 655 GCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 714
QY 541 AAAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 715 AAAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 774
QY 601 CAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 775 CAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 834
QY 661 GCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 835 GCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894
QY 721 GCGCGCGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 895 GCGCGCGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 954
QY 781 GCTTGGGAGGATCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 955 GCTTGGGAGGATCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1014
QY 841 ATCGGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 1015 ATCGGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1074
QY 901 TCGATTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 1075 TCGATTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1134
QY 961 ACCGGCGGCAACAAGAGGGGCTCTATGAAGATGCGGTGGC---AAAGACCGCTG 1017
Db 1135 GCTGGCGCGAAACCAAAAGGTTTCTATGAGATGCGCTTCCCAAAAAGAGCGTTT 1194
QY 1018 GACAAACAGATATGATGCGAGCTGGGAAGCAGCATTCGTGGCAAGCCAGTCAGCTG 1077
Db 1195 GATAATGACATATTTATTTGATCAGCTTGAGCGCAAAATTCGGCGAAAGCAGCAATGCTT 1254
QY 1078 GACGAA 1083
Db 1255 GATGAA 1260

ATTORNEY/AGENT INFORMATION:
 NAME: Carroll, Peter G.
 REGISTRATION NUMBER: 32,837
 REFERENCE/DOCKET NUMBER: CASE-02443
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 705-8410
 TELEFAX: (415) 397-8338
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1079 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 US-09-170-187-7

Query Match 55.4%; Score 619.8; DB 4; Length 1079;
 Best Local Similarity 74.9%; Pred. No. 9.4e-186;
 Matches 790; Conservative 0; Mismatches 262; Indels 3; Gaps 1;

QY 17 CAATTCACGTAGAAAATCAACACAGTATTACGTTGGAAATGTCGCCGATGCGTGTG 76
 DB 4 CTATTAGCGTCGTGATGACAAAATTTATGTGGAAATGTACCAATGTGGTGGGATG 63
 QY 77 CTCCATCGCCGTACATCCATGACCAATACGCTACGACAGAGCTGCAAGCAACGTCA 136
 DB 64 CGCTATTGCGGTCAATCAATGCAAAATCTCGACCATGATGTGAAGCGACAGTTG 123
 QY 137 ATCAATCAAGGCGCTGGAACGCGTGGCCGTATGTCGCCGTATCCGTACCGACGA 196
 DB 124 CTCAATTAATCATTTGAAGCGTGTGTCAGATATTTGTCGTATCTGTCTCAACAA 183
 QY 197 TGGACGGCGCAAGAGCTTAACACTATCAACACAGAGTTAACGTCGCGTGTGCTG 256
 DB 184 TGGATGCTGCGGAAGCTTTAAACAATTAACAACAAGTAAATGTTCCCTGTCAGAG 243
 QY 257 ACATCCACTGCGACTATGCGCTGGAAGTAGCGGAATACGCGCTGATGCTGTCG 316
 DB 244 ATATTCATTTGCACTATCGATCGCGTTAAAGTCGAGAAATGAGTGGATGTTTAC 303
 QY 317 GATTTAACCTTGGCAATATCGTAATGAGAGGCTATTCGATGCTGTTGACTGTGCG 376
 DB 304 GTATCATCTCGGAACATTTGTGTGAATGATCGCGTCCGCGTGTGTTGATTTGTCGC 363
 QY 377 GCGATTAACATTTCCGATCCGATTTGCGCTTAAACGCCGAGTCCGTGGAAGATTCG 436
 DB 364 GAGAACAAAATATTCGATTCGATTTGTTGTAATGACAGGCTCTTGAAGAAAGATTTCG 423
 QY 437 AAGAAAGTATGGGGAACGACGCGGAGCGGCTGCTGGAATCTGCCATCGCATGTTG 496
 DB 424 AAGAAAATATGCGAAGCAACGCGCAAGCCTTTGTAAGATCCGCAATTCGCTATGAG 483
 QY 497 ATCATCTGATCGCTGMACTTGATGATGAGTCAAGTACAGCGTGAAGAGCTGAGCTG 556
 DB 484 AATTTCTAGATCGTCTTAATCTTGATCAAGTTTAAAGTACGATTAAGAAACCTCCGAT 543
 QY 557 TCTCGCTGTAGTCTTATGCTTTGCTGCAAAACAGATGATCAGCCGTTGCAATCTGG 616
 DB 544 TCTTACGCGTGAATCTTAATCGTTTACTGCTTAAGCAATTAAGACGCTTTACATTTAG 603
 QY 617 GGAATCAGCGAAGCGGCTGGGCGGACGCGGCGGACAGTAAATCCGCATTTGTTAGTGC 676
 DB 604 GCAATTCAGAGAGAGGTGGGCGACGCGCTGAGTGAAGTAAATTTGAGTGGGTTTAGGAA 663
 QY 677 TGTGCTGTCTGAAGCATCGGACAGCGTGGCGTATCGCTGGGCGGCGATCCGCTG 736
 DB 664 TGTATTAGCTGAGGCGATTTGGGATACACTACGCGCTCTTTTGGGCGGCGATCCTGTAG 723
 QY 737 AAGAGATCAAAAGTGGTTTGCATATTTTGAATGCTGCGTATCCGTTCCGAGGATCA 796
 DB 724 AAGAAATCAAAAGTGGTTTGAATTTTGAATCTTACGAGATTGCTTCAAGAGGATTA 783
 QY 797 ACTTATCGCTGCGCGGACCTGTTCCGCTCAGGAATTTGATGTTATCGGTACGGTTAACG 856

DB 784 ACTTATGCTTGGCCCAACCGTTCTCGCCAGAAATTTGATGATGATACATTAATG 843
 QY 857 CGGTGGAGCAACGCGTGGAGATATCATTCATCCGATGAGAGTTCGATATGCGTGGC 916
 DB 844 CGGTAGAACAGCGCTTGAAGATATTTATTCACCAATGATGATCTATTAATCGGTTGTG 903
 QY 917 TGTGAATGGCCAGGTGAGGCGCTGTTTCTACACTCGGCGTACCCGCGCAACAGA 976
 DB 904 TAGTGAATGGTCTGGGAGGCACTGCTTCGATCTCGGCGTAAAGCGGCGGTAAACAA 963
 QY 977 AAGCGGCTCTATGAAGATGCGTGGC---AAGACGCTTGGACAAACAGATATGA 1033
 DB 964 AAGCGGCTTATTTATCTTGACGAGAGAACGCCAAAAGAGCTTTGATTAACGAATATAG 1023
 QY 1034 TCGACGAGCTGGAAGCAGCATTCGTCGAAAGCC 1068
 DB 1024 TGAACCAATTAAGAAAGCAAAATTCGTGCAAAAGTC 1058

RESULT 7
 US-08-827-190-10/C
 Sequence 10, Application US/08827190
 Patent No. 5858367
 GENERAL INFORMATION:
 APPLICANT: Rather, Philip N.
 TITLE OF INVENTION: Methods For Screening For Antimicrobials
 TITLE OF INVENTION: Utilizing aarc And Compositions Thereof
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Medlen & Carroll, LLP
 STREET: 220 Montgomery Street, Suite 2200
 CITY: San Francisco
 STATE: California
 COUNTRY: United States of America
 ZIP: 94104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/827,190
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Carroll, Peter G.
 REGISTRATION NUMBER: 32,837
 REFERENCE/DOCKET NUMBER: CASE-02443
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 705-8410
 TELEFAX: (415) 397-8338
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 886 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-08-827-190-10

Query Match 47.8%; Score 534.6; DB 2; Length 886;
 Best Local Similarity 75.3%; Pred. No. 6.7e-159;
 Matches 666; Conservative 0; Mismatches 219; Indels 0; Gaps 0;

QY 83 TCCCGCTACAGTCCATGACCAATACCGGTACGACAGACGCTGAGCAAGCAAGCTCAATCAAA 142
 DB 886 TTGCTGTCCATATATATGAGAAATATGCGGACAGATGTTGAGGCACTGTGCGGCAAA 827
 QY 143 TCAAGGCGCTGGAAGCGTGTGGCGCTGATATCGCTGATCCGTACCGAGATGAGAG 202
 DB 826 TCCATACACTTGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 767

QY 203 CGGAGAGAGCTTCAAACTCATCAACAGCAGGTTAACGTCGCGTGGTGCATCC 262
 Db 766 CAGCAGAGAGCTTAAATTAATTAAGCAGCGCGTGAATGTCATGTTGGCGATATTC 707
 QY 263 ACTTGACTATTCGATTCGCTGGAAGTAGCGGAATACGGCTGCATGTCGCTATTA 322
 Db 706 ACTTGTACATCCGATTCGCGATGAAGTGGCTGAATATGGTGTACTGCTACCAATTA 647
 QY 323 ACCGTGCAATTCGCTGATATGAAGAGCGTATTCGATGTCGCTGTCGCGCGATA 382
 Db 646 ACCGAGGATATTCGCGACTGAGAGCGTATTCGCGAGTCTGTATAGTGTCTGATC 587
 QY 383 AAAACATTCCGATTCGCTATTTGGCTTAAACGCCGATCCGTGAAAAAGATCTGCAAGAAA 442
 Db 586 ACAACATTCTTATTCGCTATGAGGGTCAATGGCGGCTCACTGAAAAAGATATCCAAA 527
 QY 443 AGTAGGGGAACCGAGCGCGGCTGCTGATATTCGCTGCTGCTATGATATC 502
 Db 526 AATACGGTGAACCAACCGTGAAGCTGTTGAATCAGCAATGCGACATGTTGATATCT 467
 QY 503 TCGATCGCTGAACTTCGATCAAGTCAAAAGTCAAGCGCTGAAAGCGTCTGACGCTTCG 562
 Db 466 TGGACAGGCTGAATTCGATCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 407
 QY 563 CTGTTGACTCTTATTCGCTTGTGCGCAAAACAGATCAGTCAAGCGTTCGATCGGAGATCA 622
 Db 406 CCGTGGCTCTTATTCGCTTATTTGCGCAAAATGATCAACACTTCACCTCGGATTA 347
 QY 623 CCGAAGCGGCTGCTGCGCGCAACCGGGGCAATTAATCCGCGCTGCTGCTGCTGCTG 682
 Db 346 CAGAAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 287
 QY 683 TGTCTGAAGGCAATTCGCGCAACGCTGCGGCTGCGCGCGCTGCGCGCGCGCGCGAGA 742
 Db 286 TGGCTGAAGGCTATTCGCGCAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 227
 QY 743 TCAAGTGGCTTTCGATATTTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 802
 Db 226 TGAAGTGGCTTTCGATATTTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 167
 QY 803 TGGCTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 862
 Db 166 TGGCTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 107
 QY 863 ACCAAGCGCTGGAATATCATCATCTCCGATGAGCTTTCGATATTCGCTGCTGCTGCTG 922
 Db 106 ACCAAGCGCTGGAATATCATCATCTCCGATGAGCTTTCGATATTCGCTGCTGCTGCTG 47
 QY 923 ATGCCCGAGTGAAGCGCTGCTTTCATCACTCGCGCTCACCGCGC 967
 Db 46 ATGCCCGAGTGAAGCGCTGCTTTCATCACTCGCGCTCACCGCGC 2

RESULT 8
 US-09-170-187-10/c
 Sequence 10, Application US/09170187
 Patent No. 6383745
 GENERAL INFORMATION:
 APPLICANT: Rather, Phillip N.
 TITLE OF INVENTION: Methods For Screening For Antimicrobials
 TITLE OF INVENTION: Utilizing aarC And Compositions Thereof
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Medlen & Carroll, LLP
 STREET: 220 Montgomery Street, Suite 2200
 CITY: San Francisco
 STATE: California
 COUNTRY: United States of America
 ZIP: 94104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/170,187
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/827,190
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Carroll, Peter G.
 REGISTRATION NUMBER: 32,837
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 705-8410
 TELEFAX: (415) 397-8338
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 886 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-09-170-187-10
 Query Match 47.8%; Score 534.6; DB 4; Length 886;
 Best Local Similarity 75.3%; Pred. No. 6.7e-159;
 Matches 666; Conservative 0; Mismatches 219; Indels 0; Gaps 0;
 QY 83 TGGCGGTACAGTCCATAGCCATAGCGGTAGCAGCAGCTGCAAGCAACGGTCATCA 142
 Db 886 TTGCTGTCCAACTTATGAGCAATACCGCGCAGCGATGTTGAACCCACTGTGCGGCA 827
 QY 143 TCAAGCGCTGGAAGCGCTGCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 202
 Db 826 TCCATATCACTTGAAGGTAGTGTGATATGTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 767
 QY 203 CGGAGAGAGCTTCAAACTCATCAACAGCAGGTTAACGCTGCTGCTGCTGCTGCTGCTGCT 262
 Db 766 CAGCAGAGAGCTTAAATTAATTAAGCAGCGCGTGAATGCTGCTGCTGCTGCTGCTGCTGCT 707
 QY 263 ACTTGACTATTCGATTCGCTGGAAGTAGCGGAATACGGCTGCATGTCGCTATTA 322
 Db 706 ACTTGTACATCCGATTCGCGATGAAGTGGCTGAATATGGTGTACTGCTACCAATTA 647
 QY 323 ACCGTGCAATTCGCTGATATGAAGAGCGTATTCGATGTCGCTGTCGCGCGATA 382
 Db 646 ACCGAGGATATTCGCGACTGAGAGCGTATTCGCGAGTCTGTATAGTGTCTGATC 587
 QY 383 AAAACATTCCGATTCGCTATTTGGCTTAAACGCCGATCCGTGAAAAAGATCTGCAAGAAA 442
 Db 586 ACAACATTCTTATTCGCTATGAGGGTCAATGGCGGCTCACTGAAAAAGATATCCAAA 527
 QY 443 AGTAGGGGAACCGAGCGCGGCTGCTGATATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 502
 Db 526 AATACGGTGAACCAACCGTGAAGCTGTTGAATCAGCAATGCGACATGTTGATATCT 467
 QY 503 TCGATCGCTGAACTTCGATCAAGTCAAAAGTCAAGCGCTGAAAGCGTCTGACGCTTCCTG 562
 Db 466 TGGACAGGCTGAATTCGATCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 407
 QY 563 CTGTTGACTCTTATTCGCTTGTGCGCAAAACAGATCAGTCAAGCGTTCGATCGGAGATCA 622
 Db 406 CCGTGGCTCTTATTCGCTTATTTGCGCAAAATGATCAACCACTTCACCTCGGATTA 347
 QY 623 CCGAAGCGGCTGCTGCGCGCAACCGGGGCAATTAATCCGCGATGTTGATAGTCTGCTG 682
 Db 346 CAGAAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 287
 QY 683 TGTCTGAAGGCAATTCGCGCAACGCTGCGGCTGCGCGCGCTGCGCGCGCGCGAGA 742
 Db 286 TGGCTGAAGGCTATTCGCGCAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 227
 QY 743 TCAAGTGGCTTTCGATATTTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 802

Db 226 TGAAGTCGGTTTGAATTTCTAAATCGTTACGATCCCTCAGCGGATCACTTTA 167
QY 803 TCGCTCCCGGACCTTTCGCGTCAAGATTGATTCGATAGCGTTAAACCGCTGG 862
Db 166 TTGCTGCCCAACCTGTTCACGCCAAGATTGATGATGGTACGGTAAATGCTTGG 107
QY 863 AGCAAGCCCTGGAAGATATCATCTCCGATGACGTTTCGATTCGCGCTGCCTGTGA 922
Db 106 AGCAGCCCTCGAAGATATATCATCTCCGATGATGATGCTCTATATTTGTTGTACTGA 47
QY 923 ATGCCCCAGGTGAGCGCTGTTCTACACTCGCGCTCACCAGCG 967
Db 46 ATGCCCCGGGTGAAGCGGAGTTTCTACTTATGCTGCGCTGGCG 2

RESULT 9

US-08-827-190-9/c
Sequence 9, Application US/08827190
Patent No. 5858367
GENERAL INFORMATION:
APPLICANT: Rather, Philip N.
TITLE OF INVENTION: Methods For Screening For Antimicrobials
TITLE OF INVENTION: Utilizing aarc And Compositions Thereof
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/827,190
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: CASE-02443
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ. ID NO.: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 886 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-827-190-9

Query Match 24.0%; Score 269; DB 2; Length 886;
Best Local Similarity 56.5%; Pred. No. 4.8e-75;
Matches 500; Conservative 0; Mismatches 385; Indels 0; Gaps 0;

QY 83 TCGCCCTACAGTCATGACCAATACCGGTCAGACAGACGTCGAAGCAACGCTAATCAAA 142
Db 886 TTGTCATCCAAAGCATGACAAACAAAACACATGATGAGCAAGCAGCGTTGCCGAAA 827
QY 143 TCAAGCGCGTGAAGCGCTGGCGTATATGTCGTTGATCCGATACGAGCATGGAGC 202
Db 826 TTACCGCTTTGGCTGAAGCGGATGCCAATCTTGGGGTAGCATATCCGATGAACGGG 767
QY 203 CGGCAAGACGTTCAAACTCATCAACACAGCGTTAAGCTGCGCGTGGTGGTACATCC 262
Db 766 CGGCAAGACGATGCGGATATATTAAGAGCGATTTCCATTCCTCTCGTGTGGATACATC 707

QY 263 ACTTGACATATCGCATTTGCCCTGAAGATAGCGATACGCGCTGATGTGCTGCGTATTA 322
Db 706 ATTTCGATTTAAACCTTGCGTTGAAGCCATTTGAAGCGCGCGGAGATTAATATCCGATCA 647
QY 323 ACCCTGCATATCGGTAATGAAGAGCGTATTCGCATGCTGTGTTGACTGTGCGCGCATTA 382
Db 646 ACCCGGCAATATCGCGCGCGCGGCAAAAAGTTGAAGCGGTGTTTAAAGCGCGCAACAA 587
QY 383 AAACATTCGATCCGATTGCGCTTAAAGCCCGGATCGCTGCGGAAAAGATCTCGCAAAA 442
Db 586 AAGCATTCGCGATCAGATGAGAGTAAACCGCGGTTTATTTGAGAAAACGATTTTGAAGA 527
QY 443 AGTATGCGAAGCAGCGCGCGAGCGCTTGTGTAATCTGCATCGCATGATGATGATC 502
Db 526 AATACGCTTATCGCACTGCCGATGAGATGATGAGAAACGACACTTACATTAATTAATTC 467
QY 503 TCGATCGCTGCACTGATGATCAGTTCAAACTGACGCTGGAAGCGTCTGACGCTCTCG 562
Db 466 TTGAGATCTTGAATTTTACGATATTTATGTCACGATGAAGGCGCTGAGCTGAACCTTG 407
QY 563 CTGTTGAGCTCTTATCGCTTGTGCTGCAAAACAGATCGATCAGCGCTGATCTGCGGATCA 622
Db 406 CATTCGAGCGTTTATGAAGAAAGCAGGAAAGCGTTTGACTACCGCTTACCTCGGATCA 347
QY 623 CCGAAGCGGTGTGCGCGAGCGGCGAGTAAATCGCCATTTGTTAGTGTGCTGCG 682
Db 346 CCGATGAGCAACACTGTTTGCAGGACAGTAAAGACGAGAGACTGCGCGCATTT 287
QY 683 TGTCTGAAGCATCGGCAACGCTGCGCTATCGCTGCGCGCGGATCGCGTGGAGAGA 742
Db 286 TAAGCAAAAGCATCGGCAACCATGCGATTTCTACATCAAGACCGCTGATGAAGAG 227
QY 743 TCAAGTCGCTTTCGATATTTGAATCGCTGCGTATCGCTGCGGAGGATCAACTTCA 802
Db 226 TAAAGTACGAAGGACCTTCTGAATCTTTCGCTTACGCTCCAAATGCTCCACGCTCA 167
QY 803 TCGCTCCCGGACCTGTTCCGTCAGGAATTTGATGATGATGATGATGATGATGATGAT 862
Db 166 TCTCATGCGCGGACTTGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 107
QY 863 AGCAAGCCCTGGAAGATATCATCTCCGATGACGTTTCGATTCGCGCTGCCTGTGA 922
Db 106 AAGATATATTTCTAAGATTAAGCGCGGATTAAGTGTGTTGCTGCGCTGCTGTA 47
QY 923 ATGCCCCAGGTGAGCGCTGTTCTACACTCGCGCTCACCAGCG 967
Db 46 AGGAGCTGTTGAAGCGAGAGAGCTGATTCGAAATCCGCGCG 2

RESULT 10

US-09-170-187-9/c
Sequence 9, Application US/09170187
Patent No. 6383745
GENERAL INFORMATION:
APPLICANT: Rather, Philip N.
TITLE OF INVENTION: Methods For Screening For Antimicrobials
TITLE OF INVENTION: Utilizing aarc And Compositions Thereof
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/170,187
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/827,190
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: CASE-02443
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 886 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-09-170-187-9

Query Match 24.0%; Score 269; DB 4; Length 886;
Best Local Similarity 56.5%; Pred. No. 4,8e-75;
Matches 500; Conservative 0; Mismatches 385; Indels 0; Gaps 0;

QY 83 TCAGCGTACGTCATGACCAATACGCGTACGACAGCGTACGACCAACGTCATCAAA 142
DB 886 TTGTCATCCAAAGCATGACACACAAAACACATGATGTAGAACAGCGTTGCGGAAA 827
QY 143 TCAAGGCGCTGAGAGCGGTGGCGTGTATGCTGATCCGATACGAGAGATGAGAG 202
DB 826 TTAAACCGTTGGCTAGAGCGGATGCCCAATGCTTGGGTAGCATGTCCGATGAACGG 767
QY 203 CGGAGAAAGCGTTCAAACTCAAAACAGCAGGTTACGTCGCGTGTGCTGACATCC 262
DB 766 CGGCAAGCGCATGCGGATATTAAGAGCGATTTCCATTCCTCGTGTGTGACATAC 707
QY 263 ACTTGACATGCGCATGCGCTGGAAGTACGCGATGCGGCTGATTTCTGCTATTA 322
DB 706 ATTTGATTTAACTTGGTTGAAAGCCATTGAAGCGCGCAATAAATCCGATCA 647
QY 323 ACCCTGCAATATCGTATATGAAAGCGTATTCGATGTTGATGCGCGGATA 382
DB 646 ACCCGGCAATTCGCGCGCGCGGCAAAAGTTGAAGCGGTTGTTAAAGCGGCAAGACA 587
QY 383 AAAACATCCGATCCGATTTGGCGTTAAAGCGGATCCGTGAAAAAAGATCTGCAAGAA 442
DB 586 AAGGATTCGATCAGAAATCGGAGTAAAGCGCGGTTCAATGAAAAAGCATTTTGAAGA 527
QY 443 AGTATGCGCAACCGACCGCGCGCGGCTGCTGAATCTGCCATGCTATGTTGATCATC 502
DB 526 AATACGTTATCCGACGCGCGATGGAATGTGTAAGAGCGCACTTCATCAATTAATAATTC 467
QY 503 TCGATCGCTGAACTGATCGATCAAAAGTCAAGGCTGAAGAGCGTCTGACGCTTCCTG 562
DB 466 TTGAGGATCTTATTTTTCAGATATATATGTGACATGAGGCTCTGACGTAACCTTG 407
QY 563 CTGTTGAGTCTTATGTTTGTGCAAAAACAGATGATCAGCGCTTGCATCTGGGGATCA 622
DB 406 CAATGAGGCTTATGAAAAAGCAGGAAAGCGTTGAGTACACCGCTCACCTCGGATCA 347
QY 623 CGGAAGCGGTGTGCGCCGACCGGCGAGTAAATCCGCCATTTGGTTAGTCTGCTGC 682
DB 346 CCGATTCAGGAACACTGTTTGGCGCACAGTAAAGAGCGCACAGCACTCGCGCATTT 287
QY 683 TGTCTGAGGCGATCGGCGACAGCTGCGGTATCGTGGCGCGCGATCCGTCGAAGAA 742
DB 286 TAAGCAAGGCAATCGGAGACACCATGCGATTTCACTAAGCCCTGTAGAGAGG 227
QY 743 TCAAGTGGTTTGAATATTTTGAATCGCTGATCCGTTGCGAGGAGATCAACTTCA 802
DB 226 TAAAGTAGCAAGGAGGAGCTTCTGAATCTTTGGGCTTAACTCCCAATGCTGCACAGCTCA 167
QY 803 TCGCTGCCGACGCTGTTGCGCTCAAGAAATTTGATGTTATCGGTACGGTTAAAGCGCTGG 862

DB 166 TCTCATGCCCGACTGTCGCGCCCTATTTGAGATTGATCTATTCAGATTCGCAATGAGTGG 107
QY 863 ACCAAGCGCTGAGATATATCATCTACCTCCGATGAGAGCTTTCATTTACGGCTCGCGGTGA 922
DB 106 AAGATATATTTCTAAGATTAAGAGCGCGCATTAAGATTGCTTTCTCGGCTCGCTGTAA 47
QY 923 ATGGCCAGGTGAGCGGCTGTTTCTACACTGCGGCTCACCGCGC 967
DB 46 ACGGACCTGTGTACAGCAGAGAACTGATATCGAATGCAATGCGCGCGC 2

RESULT 11

US-09-221-017B-309
Sequence 309, Application US/09221017B
Patent No. 644799
GENERAL INFORMATION:
APPLICANT: ROSS, Bruce C.
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
NUMBER OF SEQUENCES: 1120
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: MONROY, Gladys H.
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 309:
SEQUENCE CHARACTERISTICS:
LENGTH: 1083 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1...1083
US-09-221-017B-309

Query Match 5.0%; Score 55.4; DB 4; Length 1083;

Best Local Similarity 57.18; Pred. No. 1.5e-07;
Matches 101; Conservative 0; Mismatches 76; Indels 0; Gaps 0

QY	575	ATCGTTTGCTGGCAAAACAGATCATGATGCGCGTTCGATCTGGGGATATCAACCAACCCGGT	63
Db	281	AACCAATGGATGTGCGAGACATGACATTCGCTTCAATGGGTGTCACCGAGCGCTGGCG	34
QY	635	GTGGCGCGCAGGGGGCAGTAAATCCGCCATGGTTTAGTGTCTGCTGTCTGTGAAGCA	69
Db	341	ATGTGTGAGAGACGGCGGATCAAGAGTGTCTGTGGCATATGGTTCGCTTTTGCTGTATGGCA	40
QY	695	TGCGGACACCGCTCGCGGTATCGCTGGCGGCCCATCCGGTGTGAAGAGATCAAAAGTCG	75
Db	401	TTGGCGATGCCATTGCTGTTCACCTGTAGTAGAGATCCGAGACATGAGATATTCGGTAG	45

RESULT 12
US-09-153-599A-2/C

APPLICANT: Weber, J. Mark
 APPLICANT: Luu, B. Minh
 TITLE OF INVENTION: Method for Strain Improvement of
 TITLE OF INVENTION: Erythromycin Producing Bacterium
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:

COMPUTER READABLE FORM.
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.3C
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/153,599A

Query Match	3.68;	Score 40.6;	DB 4;	Length 1206;
Best Local Similarity	45.9%;	Pred. No. 0.0074;		
Matches 139; Conservative	0;	Mismatches 164;	Indels 0;	Gaps 0

QY	459	GCGCAGAGCCCTTCGGAAATCCATACGCATAGTGTATATCATCTGCATCGAGTCGGCTGAATC	518
Db	1120	GCGCCCTGCATCAGCTCGTACAAACAGCGTGGCCGGTGGCGGATCGGCTTGGTGAAGATCT	1061
QY	519	CGATCACTTAAAGTCAAGCTGAAACGCTGTACAGCTTCTTCGCTGTGAGCTTTATCG	578
Db	1060	GCACAGAGTACCGCTCTCTCCGCGGTGCAGCAGGATGACCGCTGCTCTTTCGACCGTTCGA	1001

QY 579 TTTGCTGGCAAAACAGATCGATCAGCCGTTGCATCTGGGGATCACCGAAGCCGCTGCTGC 638

Dy
639 GCGCAGGCGGGCAGCTAAATCCGCCATTGTTCCTTGAAGGCATCGG 6398
||| ||| ||| ||| ||| ||| ||| ||| |||
Dd 940 GCGTGCGCAGGACTCCACCCCCGGGCTCGCCCATCGCTTGATGTGGTAGAATGTCGC 881

D

QY 699 880	<div>CGACACGCTGGCCGTATCGCTGGCGGCCGATCCGGTGGAGAGATCAAGAATCGATTTCGA</div> <div> </div> <div>CGGTGGCCACGCGCATGTGCTGGCAGCGCGCGCGCGGTGAAATCCAGGACTCTCTCGA</div>	758 821
------------------	--	----------------

QY	759	TAT	761
Db	820	TCT	818

RESULT 13
US-09-153-599A-1/c
; Sequence 1, Application US/09153599A

APPLICANT: Weber, J. Mark
 APPLICANT: Lau, B. Minh
 TITLE OF INVENTION: Method for Strain Improvement of
 TITLE OF INVENTION: Erythromycin Producing Bacterium
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Rockey, Milnrow & Katz, Ltd.
 STREET: 180 N. Stetson Avenue, 2 Prudential Plaza
 CITY: Chicago
 STATE: Illinois
 COUNTRY: U.S.A.

COMPUTER READABLE FORM: `'
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.3
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/153,599A

Query Match	3.6%;	Score 40.6;	DB 4;	Length 2299;
Best Local Similarity	45.9%;			
Matches 139; Conservative	0;	Mismatch 164;	Indels 0;	Gaps 0

QY 459 GCCCAGGCGTTGCTGGAAATCTGCGCATGCGTCAATGTTGATCAATCTGCATGCGCTGAAACT 518
 Db 1432 GCCCGCTGATCAGTCTGCTGTAAGAAGCGGTGCGGCGGTGCCGATCGGCTTGGTGAAGATCT 1377
 QY 519 CGATCACTTAAAGTACGCTGGAAGAAGGTGTGACGCTTCTCGTGTGAGCTTATCG 578
 Db 1372 GCACAGAGTACCGCTCCTCGTCGGGCTGACAGAGATGCCGTGCTCTTGAAGGTCTCGA 1311
 QY 579 TTTTGCTGGCAAAACGATCGATCAGCCGTTGCATCTGGGGATCACCGCAAGCGCGTGCTGC 638

[illegible]

QY 406 GTTACCCCGGATCGCTGGAATAAGATCTGCAGAAAAAGTATGGCGAACCCGACG 465
Db 3180093 GTCAACGCCGCTGCTGGAACCGTTCAATGAGAGATATGCAACGACGCCGAG 3180034
QY 466 GCCTTGTGGAATCTGCATGCTCATGTATCATCTCATGCGCTGAACCTTGATCAG 525
Db 3180033 GCGCTGTGTGATCGCGGCTGAGGAGGCTTGCGCTTTGAGAGAGCATGGCTTCGAG 3179974
QY 526 TTCAAGTCAAGCGTGAAGGCTCTGACGCTTCTCTGCTGTGAGCTTATCGTTGCTG 585
Db 3179973 ATCAAGATCAAGCGTCAAGGATCAAGCCGCGGTGATGATGATGCGCGCTTACGAGCTGCTT 3179914
QY 586 GCAAAACAGATCGATCGGCTTCAATCTGGGATCAAGCAAGCCGAGGCTGGCGCCGACG 645
Db 3179913 GCTGCAAGCTGCGACTACGACTGCACTCGGTGCTGATCAGGAGCCGCGCTTCTTCAG 3179854
QY 646 GGGGCAAGTAAATCCGCAATGCTTATGCTCTGCTGCTGTGAAGGATCGGCGACAG 705
Db 3179853 GGCACATCAAGTCCGCTGCTGCGGCTTCTGCTGCGGCGGATAGGCGACAGC 3179794
QY 706 CTGCGCTATCGCTGCGGCGGATCCGCTGAGAGATCAAGTCTGCTGATATTTTG 765
Db 3179793 ATCCGGGTGCTGCTGCGGCGGCGGATCAAGTCAAGTGGCAATCAGGTTCTC 3179734
QY 766 AAATCGCTGCTATCCGCTGCGAGGATCACTTCACTGCGCTGCGGCTGCTGCGG 825
Db 3179733 GAGTCTTGAACCTGCGGCGGCTGCTGCTGAGATGCTGCTGCGGCTGCTGCGG 3179674
QY 826 CAGGAATTTGATGTTATCGGTACGCTTACGCGCTGAGCAACGCTGGAAGATATCATC 885
Db 3179673 GCGCAAGTCAAGCTTACGCTGCGGCAAGGATCAAGCGGCTGATGCTGCTGAT 3179614
QY 886 ACTCCGATGAGCTTTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 945
Db 3179613 GTGCGCTTCCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3179554
QY 946 TCTACACTCGGCGTCAACCGGCGGCAACAGAAAAAGCGCTTATGAAGATGGCGTCCG 1005
Db 3179553 GCGGACTTGGGCGTGGCGGCGGCAAGGATCAAGTCTTGTGAGGCGGCAAGTG 3179494
QY 1006 AAGAGCCGCTGCAACAAAGATATGATGACCACTG 1044
Db 3179493 ATCAAGACCGTGGCGGAGACAGATGTCGAGACGCTG 3179455

Search completed: May 10, 2003, 18:49:30
Job time : 7711 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2003, 16:41:08 ; Search time 197 Seconds
(Without alignments)
7058.816 Million cell updates/sec

Title: US-09-921-992-3

Perfect score: 1119

Sequence: 1 atgcataaccaggtcccaat.....ttcagcaggttgaaataa 1119

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 783854 seqs, 621352456 residues

Total number of hits satisfying chosen parameters: 1567708

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published_Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1119	100.0	1119	10	US-09-921-992-3
2	403.6	36.1	640681	10	US-09-921-992-3
3	293.4	26.2	1083	10	US-09-921-992-3
4	243.8	21.8	1164	9	US-09-921-992-3
5	239.4	21.4	1134	9	US-09-921-992-3
6	211.8	18.9	1180	10	US-09-921-992-3
7	70.6	6.3	384	10	US-09-921-992-3
8	68.4	6.1	566	10	US-09-921-992-3
9	56.8	5.1	670	10	US-09-921-992-3
10	52.6	4.7	293	10	US-09-921-992-3
11	52.6	4.7	293	10	US-09-921-992-3
12	52.6	4.7	293	10	US-09-921-992-3
13	50	4.5	403	10	US-09-921-992-3
14	48.8	4.4	564	10	US-09-921-992-3
15	48.4	4.3	379	10	US-09-921-992-3
16	48.4	4.3	528	10	US-09-921-992-3
17	48.4	4.3	2520	10	US-09-921-992-3
18	48.2	4.3	353	10	US-09-921-992-3
19	48.2	4.3	464	10	US-09-921-992-3

20	47.2	4.2	601	10	US-09-921-992-3	Sequence 11, Appl
21	45.4	4.1	613	10	US-09-921-992-3	Sequence 32, Appl
22	45.2	4.0	432	10	US-09-921-992-3	Sequence 14, Appl
23	45	4.0	988	10	US-09-921-992-3	Sequence 13, Appl
24	43.6	3.9	443	10	US-09-921-992-3	Sequence 12, Appl
25	43.4	3.9	291	10	US-09-921-992-3	Sequence 43, Appl
26	41.8	3.7	577	10	US-09-921-992-3	Sequence 28, Appl
27	41	3.7	422	10	US-09-921-992-3	Sequence 29, Appl
28	41	3.7	430	10	US-09-921-992-3	Sequence 2, Appl
29	40.2	3.6	3675	10	US-09-921-992-3	Sequence 346, Appl
30	39.6	3.5	671	9	US-10-184-634-346	Sequence 346, Appl
31	39.6	3.5	671	9	US-10-184-634-346	Sequence 346, Appl
32	38.8	3.5	388	10	US-09-921-992-3	Sequence 36, Appl
33	37.2	3.3	511	10	US-09-921-992-3	Sequence 34, Appl
34	37.2	3.3	705	10	US-09-921-992-3	Sequence 18, Appl
35	36.6	3.3	395	10	US-09-921-992-3	Sequence 45, Appl
36	35.6	3.2	211	10	US-09-921-992-3	Sequence 6, Appl
37	35	3.1	1383	9	US-09-970-989-6	Sequence 189, Appl
38	35	3.1	1383	10	US-09-967-7684-189	Sequence 612, Appl
39	35	3.1	1522	9	US-09-954-531-612	Sequence 20, Appl
40	35	3.1	1522	10	US-09-962-436-20	Sequence 26, Appl
41	34.8	3.1	3624	9	US-10-108-605-216	Sequence 9, Appl
42	34.2	3.1	1223	10	US-09-812-102-9	Sequence 589, Appl
43	33.6	3.0	425	9	US-10-025-380-589	Sequence 589, Appl
44	33.6	3.0	425	10	US-09-922-217-589	Sequence 589, Appl
45	33.6	3.0	425	10	US-09-833-263-589	Sequence 589, Appl

ALIGNMENTS

RESULT 1
US-09-921-992-3
Sequence 3, Application US/09921992
Patent No. US2002069426A1
GENERAL INFORMATION:
APPLICANT: Boronol, Albert;
APPLICANT: Campos, Narciso;
APPLICANT: Rodriguez-Concepcion, Manuel;
APPLICANT: Rohmer, Michel;
APPLICANT: Seeman, Myriam;
APPLICANT: Valentin, Henry E.;
APPLICANT: Venkatesh, Tyamagondlu V.;
APPLICANT: Venkatesh, Mylavarapu
TITLE OF INVENTION: Methyl-D-Erythritol Phosphate Pathway Genes
FILE REFERENCE: 16516.107/35-21(51897)US
CURRENT APPLICATION NUMBER: US/09/921,992
CURRENT FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: US 60/223,483
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 85
SEQ ID NO 3
LENGTH: 1119
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1119)
US-09-921-992-3
Query Match 100.0%; Score 1119; DB 10; Length 1119;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCATACCGGCTCCATTCAGTGAAGAAATCAACAGTATTTACGTTGGCAATGTG 60
DB 1 ATGCATACCGGCTCCATTCAGTGAAGAAATCAACAGTATTTACGTTGGCAATGTG 60
QY 61 CCGATTGGCGATGGTGGTCCCGATGCGGCTACAGTCAATGACCAATGACGATGACGAGAC 120
DB 61 CCGATTGGCGATGGTGGTCCCGATGCGGCTACAGTCAATGACCAATGACGATGACGAGAC 120
QY 121 GTCGAACCAAGCTCAATCAATCAATCAAGGCGCTGGACGCGTTGGCGGTGATATCGTCGT 180

121 GTGAGAGCAACGGTCAATCAATCAAGGCGCTGGAACGCGTGGCGCTGATTCGTCGCT 180
181 GTTCCGTACCGAGATGAGACGGGCGAGAACGCTTCAATCTCAACACGAGGTTAAC 240
181 GTATCCGTACCGAGATGAGACGGGCGAGAACGCTTCAATCTCAACACGAGGTTAAC 240
241 GTGCGCGTGTGCTGACATCCACTTCGACTATCGCATTCGCTGGAAGAGCGAATAC 300
241 GTGCGCGTGTGCTGACATCCACTTCGACTATCGCATTCGCTGGAAGAGCGAATAC 300
301 GCGGTGATTTGCTGCGTATTAACCTGCGCATATCGGTATATGAAGCGTATTCGATG 360
301 GCGGTGATTTGCTGCGTATTAACCTGCGCATATCGGTATATGAAGCGTATTCGATG 360
361 GTGTTGACTGTGCGGCGGCTAAACATTCGATCGCTGTTGGCGTTACGGCGGATGCG 420
361 GTGTTGACTGTGCGGCGGCTAAACATTCGATCGCTGTTGGCGTTACGGCGGATGCG 420
421 CTGGAAGAAAGATCTGCAAGAAAGATGCGGAACCGACGCGGAGCGGCTTGAATCT 480
421 CTGGAAGAAAGATCTGCAAGAAAGATGCGGAACCGACGCGGAGCGGCTTGAATCT 480
481 GCGATCGCTATGTTGATCATCTCGATCGCTGATCGCTGATCGATCGATCGATCGAT 540
481 GCGATCGCTATGTTGATCATCTCGATCGCTGATCGCTGATCGATCGATCGATCGAT 540
541 AAGCGCTGACGCTCTCCGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 600
541 AAGCGCTGACGCTCTCCGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 600
601 CAGCGCTGATCTGAGGAGATCAGCGAAGCGGCTGCTGCGGAGCGGAGGATGATGAT 660
601 CAGCGCTGATCTGAGGAGATCAGCGAAGCGGCTGCTGCGGAGCGGAGGATGATGAT 660
601 CAGCGCTGATCTGAGGAGATCAGCGAAGCGGCTGCTGCGGAGCGGAGGATGATGAT 660
661 GCGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
661 GCGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
721 GCGCGCGATCCGCTGCAAGAGATCAAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
721 GCGCGCGATCCGCTGCAAGAGATCAAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
721 GCGCGCGATCCGCTGCAAGAGATCAAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
781 CGTTCGAGAGGATCAACTTCATCGCTGCGCAGCTGCTGCTGCTGCTGCTGCTGCTG 840
781 CGTTCGAGAGGATCAACTTCATCGCTGCGCAGCTGCTGCTGCTGCTGCTGCTGCTG 840
781 CGTTCGAGAGGATCAACTTCATCGCTGCGCAGCTGCTGCTGCTGCTGCTGCTGCTG 840
841 ATGCGTACGCTTAACGCGCTGAGCAACGCTGGAAGATATCACTCCGATGAGCTG 900
841 ATGCGTACGCTTAACGCGCTGAGCAACGCTGGAAGATATCACTCCGATGAGCTG 900
901 TCGATTATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
901 TCGATTATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
901 TCGATTATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
961 ACCGCGGCGCAAGAAAGCGGCTCTATGAAGATGCGCTGCGCAAGACCGCTGCGAC 1020
961 ACCGCGGCGCAAGAAAGCGGCTCTATGAAGATGCGCTGCGCAAGACCGCTGCGAC 1020
1021 AACACGATATGATGACGACGCTGGAAGCAAGATTCGTCGAAAGCGCATCGAGCTG 1080
1021 AACACGATATGATGACGACGCTGGAAGCAAGATTCGTCGAAAGCGCATCGAGCTG 1080
1081 GAAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1119
1081 GAAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1119

RESULT 2
US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. US20020127687A1
; GENERAL INFORMATION:

APPLICANT: SHIGENOBU, SHUJI
APPLICANT: WATANABE, HIDEIMI
APPLICANT: HATTORI, MASAHIRA
APPLICANT: SAKAKI, YOSHIYUKI
TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
CURRENT APPLICATION NUMBER: US/09/790,988
CURRENT FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: JP2000-107160
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 1
LENGTH: 640681
TYPE: DNA
ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 36.1%; Score 403.6; DB 10; Length 640681;
Best Local Similarity 62.2%; Pred. No. 2,9e-127;
Matches 653; Conservative 0; Mismatches 394; Indels 3; Gaps 1;

19 ATCAACGTGAAATCAACAGATATTTAGCTGGGATGTCGCGATGGCGATGGGCT 78
314293 ATCAATAGAGAAATCTGATGCTATTTATGTTGAAAGTCCCTATTTGCAATATGCG 314352
79 CCCATGCGCTGACAGTCCATGACCAATACGCTAGCAGACAGCTGCAAGCAAGCTCAAT 138
314353 CCATATGCACTTCAATCTAGTCAAAATACCTGTAATTAATGCTGTAATTAAT 314412
139 CAATCAAGGCGCTGAGACCGCTGGCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 198
314413 CAATCTTACAGTTTACAAAAGTAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTG 314472
199 GAGCGGAGAGAGCTTCAACATCAACACAGAGATTAACGCTGCTGCTGCTGCTGCTGCTG 258
314473 AAGCTGACAGATCATTTCAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 314532
259 ATCACTTGCATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 318
314533 ATCACTTGCATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 314592
319 ATTAACCTGCGCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 378
314593 ATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 314652
379 GATTAATACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 438
314653 GATTAATACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 314712
439 GAAAGTATGCGCAACCGACGCGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 498
314713 AAAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 314772
499 CATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 558
314773 TACTTGTATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 314832
559 CTGCGTGTGAGTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 314892
314833 TTAGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 314952
619 ATCAACGAAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 678
314893 ATCACTGAAATCCGCTGCTTAAAGATGAGCAAGTAAATCACTATTAATTAATTAAT 314952
679 CTGCTGCTGCAAGGCAATGCGGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 738
314953 TTATTTATGAAAGGCAATGAGGATTAATTAATTAATTAATTAATTAATTAATTAAT 315012
739 GAGATCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 798
315013 GAAGTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 315072

QY 799 TTCAATGCGCTGCCGACCTGTTCCGCTGAGAAATTTGATGTTATGCTAGCGGTTAAACGG 858
Db 315073 TTTATTTGCTTGTCTACTGTTCTTCTAGACAAATTTGATGTTATTAATACAGTAATCA 315132
QY 859 CTGGAGCAACGCGCTGGAAGATATCATCACTCCGATGAGCGTTTGCATTATCGCGTCCGT 918
Db 315133 CTAGAAAAAAATCTAGAAAGATATCTGACCTCCATGATGATCAATTAATGTTGCGCTT 315192
QY 919 GTGATGCGCCGAGGAGGCGCGCTGTTCTCTACCTCGCGGTCACCGGCGCAACAAGAA 978
Db 315193 GTTAATGGAATATGATGATCTTAATAATACCACTTTAGGCTTAGCAGAGATCATAGAAA 315252
QY 979 AGCGGCTCTATAGATGCGGT---GCGCAAGACCGCTGAGACAGCATATGATC 1035
Db 315253 AGTGCAATTTATGAGAGCGAGGTAGACAAAAGAAAAAATTAAGAAAGAAATTAATA 315312
QY 1036 GACCAAGCTGGAAGACGCGATTCGTCGAAA 1065
Db 315313 GAAAAAATGAAATTAATAATTCGAAAAAA 315342

RESULT 3

US-09-974-300-1692
Sequence 1692, Application US/09974300
Patent No. US20020146721A1
GENERAL INFORMATION:
APPLICANT: Berka, Randy M.
APPLICANT: Clausen, Ib Groth
TITLE OF INVENTION: Methods For Monitoring Multiple Gene
FILE REFERENCE: 10085,500-US
CURRENT FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: 09/680,598
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/279,526
PRIOR FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 8481
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1692
LENGTH: 1083
TYPE: DNA
ORGANISM: Bacillus licheniformis
US-09-974-300-1692

Query Match 26.2% Score 293.4; DB 10; Length 1083;
Best Local Similarity 56.6% Pred. NO. 1.2e-90;

Matches 543; Conservative 0; Mismatches 416; Indels 0; Gaps 0;

QY 22 CAACGTAGAAATCAACAGTATTTACGTTGGGATGTGCCGATTGGCGATGTCCTCC 81
Db 21 CATGCTACAAAACCGCTCCGTTAAAGTGGACCTTTAACAATAGCGCGCAATACGA 80
QY 82 ATCGCGCTAGCTCATCAACCAATACGCGTACAGACAGCTGGAAGCGGTCAATCA 141
Db 81 GTCGCTATTTCAAGCATCAACAACAACACATGACGTGAGCAACCGTCCCGAA 140
QY 142 ATCAAGGGCGTGAAGCGGTTGGCGCTGATATCGTCCGTTATCGTACCGCAGTGGAC 201
Db 141 ATCAACAGACTGCGGAGACAGGATGTCAAAATGTCGCGCTGCTGCTGATGAACGG 200
QY 202 GCGGAGAGCGTTCAACATCAACAGCAGGTTAAACGTCGCGCTGGTGGCTACATC 261
Db 201 GCTGCGAGCGCATTCAGAGATCAAAAAGGAGATATCATCCCTTGTGCTGATAT 260
QY 262 CACTTCGACTATCGATTGCGGTGAAGTACGGGAATACGGCGTGGATTGTCTGGTAT 321
Db 261 CATTTCAACTATTAATTTGCAATTAAGACGATCGAAGCGCGCATTAATAATCCGATC 320
QY 322 AACCTGCAATATGTGATATGAGAGCGTATTCGATGTTGGTGTGACGTGGCGCAT 381
Db 321 AATCGGGAATACATGCGCGCGCGGAAAAAGTTGAAGCGGTGCTCAACGCAAGGA 380

QY 382 AAAACATTCGATCCGATATGCTTAACGCCGATCGCTGGAAAAAGATCTGCAAGA 441
Db 381 AAGGCGATTTCCGATCCGATGCGGCTCAATGCAAGCTCTCTGAAAAACGATCCCTTG 440
QY 442 AAGTATGCGAAGCAGCGCGGAGCGCTTGTGGAATCTGCGATGCTATCTTATATC 501
Db 441 AAGTATGCGAAGCAGCGCGGATGCTGGAAGCGCGCTGATCAATTAATTAAT 500
QY 502 CTGATGCGCTGATCTGATGATCAAGTCAAGCAGCGGTAAGCGCTGATGCTTCTC 561
Db 501 CTGAGGATCTGATTTCCAGATATCATGTCACATGAAGCGCTGATGATTAACCTG 560
QY 562 GCTTTGAGTCTTATCGTTTGTGCAAAACAGATCATCAGCTTGGATCTGAGATC 621
Db 561 GCGATTTAGGCAATGAAAGAGCGCTAAGCTTTCGATTAATCCCTTCATTTAGCANT 620
QY 622 ACCGAAGCGGCTGTGCGCGCAGCGGCGAGTAATAATCCGATTTAGCTCTGTG 681
Db 621 ACCGAATCGGGAACGCTGTTGCGGTACAGTGAAGAGCGCGGCTCTCGCGCATC 680
QY 682 CTGCTGAAGGATGCGGACACGCGCGCTGCTGCGCGCGATCCGATCCGTAAGAG 741
Db 681 CTTTAAAGGATCGCAATACATTCGCGATTTCTTAAGCGCGGACCGCTCGAAGA 740
QY 742 ATCAAGTCTGTTGATATTTGAATCGCTGCTATCCGTCGAGAGATCAACTTC 801
Db 741 GTAAAGTCTGCGAGGAGCTCTGAATCTTTGCGGCTGCTCCAAATCGGCAACTTG 800
QY 802 ATGCGCTGCGCGACCTGTTGCGGTGAGCAATTTGATGTTATGCTAGCGTTAACGCGTG 861
Db 801 ATTCTCTCCGCACTTGGCGCGGATCGAAATCTGATTTGATGCTCAATGAATTC 860
QY 862 GAGCAAGCGCTGGAAGATATCATCTCCGATGAGAGCTTTCATATGCGCGCTG 921
Db 861 GAGATTTACATCGCAAAATCAAGCGCGGATCAAGTGGTCTGCTGCGTGGCGCTC 920
QY 922 AATGCCAGATGAGCGGTGTTCTACACTCGGCGTCAACCGGCGCAAGAAAG 980
Db 921 AAGGTCGCGGAGAAAGCGGCGGCAACCGATATTCGATGCGGCGGCGGATGATG 979

RESULT 4

US-09-712-363-110
Sequence 110, Application US/09712363
Patent No. US20020164588A1
GENERAL INFORMATION:

APPLICANT: Eisenberg, David
APPLICANT: Rotstein, Sergio H.
APPLICANT: Marcotte, Edward M.
TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
FILE REFERENCE: 07419-032001
CURRENT FILING DATE: US/09/712,363
PRIOR APPLICATION NUMBER: PCT/US00/02246
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: 60/179,531
PRIOR FILING DATE: 2000-02-01
PRIOR APPLICATION NUMBER: 60/117,844
PRIOR FILING DATE: 1999-01-29
PRIOR APPLICATION NUMBER: 60/118,206
PRIOR FILING DATE: 1999-02-01
PRIOR APPLICATION NUMBER: 60/126,593
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 60/134,093
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/134,092
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/165,124
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/165,086
PRIOR FILING DATE: 1999-11-12

; NUMBER OF SEQ ID NOS: 292
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 110
 ; LENGTH: 1164
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium tuberculosis
 ; US-09-712-363-110

Query Match 21.8%; Score 243.8; DB 9; Length 1164;
 Best Local Similarity 53.5%; Pred. No. 1.7e-73;
 Matches 534; Conservative 0; Mismatches 462; Indels 3; Gaps 1;

QY 49 GTTGGGAATGTCGGATGTCGATGTCCTCCATCGCCCTACAGTCCATGACCAATAG 108
 Db 76 GTGGGCACTCGCGCGGAGTGAACCAATCGCTCGTGAATGATGACCAACC 135
 QY 109 CGTACGACAGCGTGCAGGAGGATCAATCAATCAAGGCGTGSAGCCGTTGGCGCT 168
 Db 136 AAAACCCAGACGCTCACTGACATTCGAACAAATCGCCAGCTGACCGCGCGATGC 195
 QY 169 GATATGCTCGGTATTCGATCCGAGATGAGCGGAGCAAGCGTTCAAACTCATAAA 228
 Db 196 GATATGCTCGGTATTCGATCCGAGATGAGCGGAGCAAGCGCGTGGCGAGATGCC 255
 QY 229 CAGCAGGTTAAGCTGCGCGTGTGCTGACATCCACTTGCATATGCAATGCGCTGAAA 288
 Db 256 CGGACAGAGCGAGATCGCGGAGTGCAGGACATTCATTCAGCGCGCTCATATTCGCC 315
 QY 289 GTAGCGGAATAGGCGGCTGATGTCGCTATTAACCCCTGCAATATC---GGTAATGAA 345
 Db 316 GGCATGACCGCTGATGTCGCGGAGTGCAGGAGTCAACCCGCGCAACAGAGATTGAC 375
 QY 346 GAGCGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 405
 Db 376 GGCAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 435
 QY 406 GTTAAAGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 465
 Db 436 GTTAAAGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 495
 QY 466 GCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 525
 Db 496 GCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 555
 QY 526 TTCAAAGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGT 585
 Db 556 ATCAAGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGT 615
 QY 586 GCAAAACAGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGT 645
 Db 616 GCTGACAGGAGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGT 675
 QY 646 GGGGAGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGT 705
 Db 676 GGGGAGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGT 735
 QY 706 GTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 765
 Db 736 ATCGGAGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGT 795
 QY 766 AATTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 825
 Db 796 GATTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 855
 QY 826 CAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 885
 Db 856 GCGGAGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGT 915
 QY 886 ACTCCAGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGT 945
 Db 916 GTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 975
 QY 946 TCTACACTGCGGCTGACCGGCGGCAACAGAAAAAGCGGCTCTATGATGATGCGTGGC 1005

Db 976 GCCGACCTGGCGCTGCGCTCGGCAACGCAAGTCAATCTTTGATCGGGGGAAGTG 1035
 QY 1006 AAAGACCGTCTGACACAAACAGATATGATGACACGCTG 1044
 Db 1036 ATCAAGACCGTGGCGGAGCAACAGATGCTGAGAGCGTG 1074

RESULT 5

US-09-738-626-2205
 ; Sequence 2205, Application US/09738626
 ; Publication No. US20020197605A1

; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAOKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 2205
 ; LENGTH: 1134
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 ; US-09-738-626-2205

Query Match 21.4%; Score 239.4; DB 9; Length 1134;
 Best Local Similarity 54.1%; Pred. No. 5.6e-72;
 Matches 533; Conservative 0; Mismatches 446; Indels 6; Gaps 2;

QY 22 CAAGGTGAATCAACACGATATTAAGTGGGATGTCGCGATTTGGCGATGCTGCC 81
 Db 10 CCAAGGTGAATCAACACGATATTAAGTGGGATGTCGCGATTTGGCGATGCTGCC 69
 QY 82 ATGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 141
 Db 70 ATTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 129
 QY 142 ATCAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 201
 Db 130 ATCAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 189
 QY 202 GCGGAGAGCGTGAACACTATCAACACAGAGTGAAGTGGCGGTGGTGGATGATC 261
 Db 190 GATGCGGAGAGCGTGAACACTATCAACACAGAGTGAAGTGGCGGTGGTGGATGATC 249
 QY 262 CACTTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 321
 Db 250 CACTTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 309
 QY 322 AACCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 378
 Db 310 AACCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 369
 QY 379 GATTAATAAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 435
 Db 370 GATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 429

QY 436 CAAGAAATATGCGAACCGACGCCGCTGCTGGAATCTGCCATGCGTCACTT 495
DB 430 GCAAAATACACGAGCAACCCAGAGAGCTCTGCGAATCCGCAATGCGGAGGCC 489
QY 496 GATCATCTGCGATCGCTGAACTTGCATGCTCAAGAGGAGGAGGAGGAGGAGG 555
DB 490 GGCCTGTTTGAAGAGACAGCGCTTCGCGCATGCGCATCTGCTGAGACACTCCGACCA 549
QY 556 TTCTCGCTGTTGAGTCTTATGCTTGTGCGCAAAACAGATGACGCGTGTGATCG 615
DB 550 GTACATCATGCTGAGAGCCATCCAGCCAGCTCGCTGAACAAAGGAGACTACCCACTGCACCTC 609
QY 616 GGGATACCGAAGCCGCTGCTGCGCGACGCGGAGCAATTAATCCGCAATTTAGT 675
DB 610 GGTGTACTGAAAGCTGCTGCCAAGTTTCATGAGGAGCAATCACTTCTTACATTCGCGC 669
QY 676 CTGCTGCTGCTGAAAGGAGCTGCGGACAGCTGCGGTATGCTGCGCGCGCATCCGCTC 735
DB 670 GCTCTGCTGCTCCAGAGGACTCGCGGACACTATCCGTCTCTCTTCTGCTGACCCAGCTG 729
QY 736 GAGAGATCAAAAGTGGTTTGCATATTTTGAATGCGCTATCCGTTCCGAGAGGATC 795
DB 730 GAAGAATATCAAGGTGGGAGGACAGATTTGCAAGTCCCTCAACCTGGGCCAGCAAGCTG 789
QY 796 AACTTCATGCTGCTGCGCGACCTGCTGCGCTCAAGAAATTTGATGTTATCGCTACGCTTAA 855
DB 790 GAATTCGCTGCTGCTGCGCATGATGCGCGGACAGCTGATGATGATGATGATGATGATGAT 849
QY 856 GCGCTGAGACACGCTGGAATATATCATCATGCGATGAGAGGATTTGATTTGCGCTGTC 915
DB 850 GAAGTACCGAAGCACTGACGCGCATGGAATTTCCACTGCGCTGCTGCTGCTGCTGCTGCTG 909
QY 916 GTGTGATGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 975
DB 910 GTTGTAAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 969
QY 976 AAAAGGCGCTCTATGAAAGATGGCG 1000
DB 970 AAGGCGAGATCTTTGCAAGGCG 994

RESULT 6
US-09-881-752A-195
; Sequence 195, Application US/09881752A
; Patent No. US20020115078A1
; GENERAL INFORMATION:
; APPLICANT: Kleantous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020115078A1 Helicobacter Polypeptides in the
; FILE REFERENCE: 06132/041002
; CURRENT APPLICATION NUMBER: US/09/881,752A
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 08/833,457
; PRIOR FILING DATE: 1997-04-01
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 195
; LENGTH: 1180
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (51)...(1127)
US-09-881-752A-195

Query Match 18.9%; Score 211.8; DB 10; Length 1180;
Best Local Similarity 52.0%; Pred No. 2e-62;
Matches 524; Conservative 0; Mismatches 477; Indels 6; Gaps 2;

QY 31 AATCAACAGATATTTACGTTGGGAATGTCGCCGATTTGGGATGTCCTCCATCGCCGTA 90
DB 69 AAGACCAAGCAATTTTATGCTGGGCGGCGCATATGAGGGGATGATGCTCCATATAGCAG 128
QY 91 CAGTCATGACCAATATGCGGTACGACAGCTGCAAGCAAGCGTCAATCAATCAAGCGG 150
DB 129 CAAGCATGACCTTTAGCAAAACCGCTGATTTGAAGCACTTAAATCAATTTGACAGA 188
QY 151 CTGGAAGCGGTTGGGCGTATGTCGCTGATCGGTACGACATGAGAGCGGCGAGGA 210
DB 189 CTCAAACTGCGCGGCGGCGATTTAGTGAAGGCGGAGTGAATGAAGAAAGACGCTCTA 248
QY 211 GCGTTCAACATCATCAACAGCAGGTTAAGCTGCGCGCTGCTGATCCATCCATCGAC 270
DB 249 GCGTTAAAGAAATTTGAAGAAAGATGTCCTTGGCTTTATGCTGATATTTTCATTTCCAT 308
QY 271 TATCGCATGCGCTGAAGTACCGGAAATAGCGGCTGATGTTCTGCTGATTAACCTGCG 330
DB 309 TATTAATTCGCTCTCATGTCCTCATAA--CGGTGATGCGATCAGGATTAACCCCGGA 365
QY 331 AATATCGTAATGAAGAGCGTATTCGATGAGGTTGACTGTCGCGCGGATTAAGAAAT 390
DB 366 AACATCGGCTTAAGAGAGAGATCAAGCGGCTGATGCTTTGTAAGCAAAAGAAACAT 425
QY 391 CCGATCCGTAATTTGGGCTTAACCGCGGATCGTGAAGAAAGATCTGCAAGAAAGATAGCG 450
DB 426 CCTATAAGAAATTTGGGCTGAATGCTGGAGATTTAGAAAGACATTTGATCAAAATA 482
QY 451 GAACGAGCGCGCAGCGCTGCTGGAATGCTGCAATGCTGATGCTGATGATGATGATGATGAT 510
DB 483 GAGCCACCCCAAAAGGCAATGTAAGAACGCTTTGTAAGAACGCAAACTTTTGAAGAT 542
QY 511 CTGAATTCGATCAATCAATCAAGCTGAAGGAGTGTGAGCGTCTTCTGCTGCTGAG 570
DB 543 TTGATTTTACCAATTTTAAGATTTCTTTAAAGAGAGGATGATTTGCAACATAGAA 602
QY 571 TCTTATCGTTTCTGCGCAAAACAGATGATGATGATGATGATGATGATGATGATGATGAT 630
DB 603 GCTTACAGAGATGCTTCCGCTCTGATGATCTTCTTCAATTTGGGCTTACGAGCGG 662
QY 631 GGTGGCGCGGAGGCGGCGCATTAATCCGCTATGTTAGTCTGCTGCTGCTGCTGCTG 690
DB 663 GGAATCTTTTATGCTCCAGTATCAATCCGCTATGCTTATGAGGGGCTTTTAAATGAG 722
QY 691 GGCATCGCGACACGCTGCGCTATGCTGCGCGCGATCCGCTGAAAGATCAAAAGTC 750
DB 723 GGCATTTGGGATACGATGCGCTATCCATCAACAGGAGATTTGAAGAAATCAAAAGTC 782
QY 751 GGTTCGATATTTTGAATGCTGCTGATGCTGCTGCGAGGAGATCAATTCATCGCTGCG 810
DB 783 GCCAGAGCAATTTTACGCTATGCGCATAGCGGCGGTTGAAGAAAGGAGATTAATGATTTCTG 842
QY 811 CGACCTGTTCCGCTGAGCAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 870
DB 843 CCACCTGCGGCGATGTAAGCAATTTATGATGATGATGATGATGATGATGATGATGATGAT 902
QY 871 CTGAAGATATCATCTGCTGATGAGAGCTTATGATGATGATGATGATGATGATGATGATGAT 930
DB 903 TTAAGCAGATCAAAACCTTTAGACATTTACGCTGATGATGATGATGATGATGATGATGAT 962
QY 931 GGTGAGGCGCTGTTTCACTGCGCTGACCGCGGCGCAACAGAAAGAGCGCTCTAT 990
DB 963 GGTGAAGCCAGCATGACATGCGATGCGCTTTTGGGAATCGAGCGGTTTATCAT 1022
QY 991 GAAGATGCGCTGCGCAAGAGCGCTGCGCAACAGATATGATGA 1037
DB 1023 AAGAGGATTAATTCACAACTGCTGAAGGATTTATTTGA 1069

RESULT 7
US-09-921-992-21/C
; Sequence 21, Application US/09921992

Db 403 GGTTCGCGAAACCATGGTGAATATGTGGCCTGTCTTCCTGTGCGC 4522

RESULT 10
US-09-974-300-6102

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	5.1%;	Score 56.8;	DB 10;	Length 348;
Best Local Similarity	54.9%;	Pred. NO. 4.2e-09;		
Matches 112; Conservative	0;	Mismatches 92;	Indels 0;	Gaps 0.

[illegible]

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

US-09-921-992-25

Query Match	4.7%	Score 52.6;	DB 10;	Length 293;
Best Local Similarity	53.7%	Pred. No. 1.1e-07;		
Matches 109;	Conservative 0;	Mismatches 94;	Indels 0;	Gaps 0;

QY 85 GCCCTCAGTCCATGACCAATAGCGGTACGACAGACGTCCGAAACGCGTCATCAATC 144
 Db 70 AGGATTCAAACCATGACACACTTCAGATACCAAGGATGTGTGGAAACAGTAGAGAGGTG 129
 QY 145 AAGGCGCTGCGAAGCGGTTGGCGCGATATATCGTCCGTATCCGTAACGACAGATGGACGG 204
 Db 130 ATGAGAGATAGCAGATTAAGGAGCTGATCTTTTGAATTAACAGTCCAGGCTAGSAAAGAA 188
 QY 205 GCAGAGCGCTCAAACTCATCA 227
 Db 190 GCTGATGCTGCTTGAGATCA 212

```

RESULT 12
US-09-921-992-26
; Sequence 26, Application US/09921992
; Patent No. US20020069426A1
;
GENERAL INFORMATION:
; APPLICANT: Boronat, Albert;
; APPLICANT: Campos, Narciso;
; APPLICANT: Rodriguez-Concepcion, Manuel;
; APPLICANT: Rohmer, Michel;
; APPLICANT: Seeman, Myriam;
; APPLICANT: Valentini, Henry E.;
; APPLICANT: Venkatesh, Tyanaagondlu V.;
; APPLICANT: Venkatesmesh, Mylavaraipu
; TITLE OF INVENTION: Methyl-D-Erythritol Phosphate Pathway Genes
; FILE REFERENCE: 16516.107/35-21(51897)US
; CURRENT APPLICATION NUMBER: US/09/921,992
; PRIOR APPLICATION NUMBER: US 60/223,483
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 85
;
SEQ ID NO 26
;
; LENGTH: 456
;
; TYPE: DNA
;
; ORGANISM: Zea mays
US-09-921-992-26

```

Query Match	4.78;	Score 52.6;	DB 10;	Length 456;
Best Local Similarity	53.78;	Pred. No. 1.4e-07;		
Matches 109; Conservative	0;	Mismatches 94;	Indels 0;	Gaps 0;

Oy 25 CGTAGAAATCTCAACCGATTTACTCTTGGGATGTGCCGATTTGGCGATGTGGTCCCATC 8
 Db 10 CGAGAGAAATCTCGAACTGTATGTTGGGGAATGTGCCACTTTGGCAGTGAATCACCATA 69
 Oy 85 GCCGTACAGTCATGACCAATACGCGTACGACAGCTGCAAGCAACGGTCAATCAATC 144
 Db 70 AGGATTCAACACCATGACGACTTCAGATTACCAAGATGTTGCGAAACACAGTGAAGAGGTG 129
 Oy 145 AAGGCGCTGGAACGGCTTGGCGCTATATGTCGCTGATTCCTGATCCGACAGATGGACGG 204
 Db 130 ATGAGGATTGCAGATTAAGGAGCTGATCTTTGTAAATAATACAGTCCAGGGGTAGGAGGAA 189
 Oy 205 GCAGAGCGTTCAACTCA 227
 Db 190 GCTGATGCTGCTTTGAGATCA 212

RESULT 13
US-09-921-992-24
; Sequence 24, Application US/09921992

QY	598	GATACCGCTTCATCTCATTGGGATATACGAGACCGGTGTGTGTGCGCGCAGCGGGCAGTAAA	657
	215	GATTATCCCTTTGCATTCTTGGGAGTTACTCGAGGACGAGGAAGGCAAGATGTGACGCGATGAA	274
Db	658	TCGCGCATGGTTTGTAGTCTCTGCTGTCTGGAAGGCAATGGGGACACCGTGTGCGTATCG	712
QY	275	TCGCGCATTTGGATTGGGACGCTTTCTTCAGGACGGGCTGTGATACCAATAAGAGTTTCA	334
Db	718	CTGGCGCGCGATCCGCGTGAAGAGAT	743
QY	335	CTGACGAGCACCAAGAAAGAGAT	360

Mon May 12 08:44:42 2003

Search completed: May 10, 2003, 17:56:02
Job time : 830 secs

us-09-921-992-3.rnpb

